

# QSite 4.0

## User Manual

Copyright © 2006 Schrödinger, LLC. All rights reserved. CombiGlide, Epik, Glide, Impact, Jaguar, Liaison, LigPrep, Maestro, Phase, Prime, QikProp, QikFit, QikSim, QSite, SiteMap, and Strike are trademarks of Schrödinger, LLC.

Schrödinger and MacroModel are registered trademarks of Schrödinger, LLC.

Python is a copyrighted work of the Python Software Foundation. All rights reserved.

The C and C++ libraries for parsing PDB records are a copyrighted work (1989) of the Regents of the University of California. All rights reserved.

See the [Copyright Notices](#) for full copyright details.

To the maximum extent permitted by applicable law, this publication is provided “as is” without warranty of any kind. This publication may contain trademarks of other companies.

Please note that any third party programs (“Third Party Programs”) or third party Web sites (“Linked Sites”) referred to in this document may be subject to third party license agreements and fees. Schrödinger, LLC and its affiliates have no responsibility or liability, directly or indirectly, for the Third Party Programs or for the Linked Sites or for any damage or loss alleged to be caused by or in connection with use of or reliance thereon. Any warranties that we make regarding our own products and services do not apply to the Third Party Programs or Linked Sites, or to the interaction between, or interoperability of, our products and services and the Third Party Programs. Referrals and links to Third Party Programs and Linked Sites do not constitute an endorsement of such Third Party Programs or Linked Sites.

Revision A, April 2006

# Contents

Document Conventions .....	vii
Chapter 1: Introduction .....	1
1.1 About QSite .....	1
1.2 Citing QSite in Publications .....	1
Chapter 2: Introduction to Maestro.....	3
2.1 General Interface Behavior .....	3
2.2 Starting Maestro .....	3
2.3 The Maestro Main Window .....	4
2.3.1 The Menu Bar .....	6
2.3.2 The Toolbar .....	7
2.3.3 Mouse Functions in the Workspace.....	10
2.3.4 Shortcut Key Combinations .....	11
2.4 Maestro Projects.....	11
2.4.1 The Project Table Toolbar .....	13
2.4.2 The Project Table Menus .....	14
2.4.3 Selecting Entries .....	15
2.4.4 Including Entries in the Workspace.....	16
2.4.5 Mouse Functions in the Project Table .....	16
2.4.6 Project Table Shortcut Keys .....	17
2.5 Building a Structure.....	18
2.5.1 Placing and Connecting Fragments .....	18
2.5.2 Adjusting Properties.....	20
2.5.3 The Build Panel Toolbar .....	20
2.6 Selecting Atoms .....	21
2.6.1 Toolbar Buttons .....	21
2.6.2 Picking Tools .....	22
2.6.3 The Atom Selection Dialog Box .....	23

<b>2.7 Scripting in Maestro</b>	23
2.7.1 Python Scripts	23
2.7.2 Command Scripts	24
2.7.3 Macros	25
<b>2.8 Specifying a Maestro Working Directory</b>	25
<b>2.9 Undoing an Operation</b>	26
<b>2.10 Running and Monitoring Jobs</b>	26
<b>2.11 Getting Help</b>	28
<b>2.12 Ending a Maestro Session</b>	28
<b>Chapter 3: QSite Tutorial</b>	<b>29</b>
<b>3.1 Preparing a Working Directory</b>	29
<b>3.2 Starting Maestro</b>	30
<b>3.3 Importing the Complex</b>	30
<b>3.4 Selecting the QSite Job Type</b>	30
<b>3.5 Defining a QM Region</b>	31
<b>3.6 Running the QSite Job</b>	34
<b>3.7 Examining Results</b>	35
3.7.1 Comparing Input and Output Structures	35
3.7.2 Comparing Ligand-Receptor Interactions	35
<b>Chapter 4: Protein Preparation</b>	<b>39</b>
<b>4.1 The Protein Preparation Facility</b>	39
<b>4.2 Step-by-Step Overview</b>	40
<b>4.3 Importing the Protein Complex Structure</b>	41
<b>4.4 Deleting Unwanted Waters</b>	42
<b>4.5 Simplifying the Protein Complex</b>	42
4.5.1 Determining Whether the Complex Is a Multimer	42
4.5.2 Retaining Needed Subunits	43

---

<b>4.6</b>	<b>Adjusting the Protein, Metal Ions, and Cofactors</b>	44
4.6.1	Checking the Protein Structure for Metal Ions and Cofactors	45
4.6.2	Deleting Protein-Metal Bonds	45
4.6.3	Adjusting Metal Ion Charges	46
4.6.4	Displaying and Adjusting the Cofactor	47
<b>4.7</b>	<b>Adjusting the Ligand</b>	48
4.7.1	Manually Deleting Protein-Ligand Bonds	48
4.7.2	Adjusting Ligand Atom and Bond Properties	49
<b>4.8</b>	<b>Running Protein Preparation on the Structures</b>	49
4.8.1	Defining the Ligand	49
4.8.2	Choosing a Procedure	50
4.8.3	Other Options	51
4.8.4	Starting the Protein Preparation Job	51
4.8.5	Output Job Files	51
<b>4.9</b>	<b>Checking the Output Structures</b>	52
4.9.1	Checking the Orientation of Water Molecules	52
4.9.2	Resolving H-Bonding Conflicts	53
<b>4.10</b>	<b>Command-Line Protein Preparation</b>	53
4.10.1	Usage Summary for <code>protprep</code>	53
4.10.2	Usage Summary for <code>pprep</code>	55
4.10.3	Usage Summary for <code>impref</code>	55
<b>Chapter 5: Running QSite From Maestro</b>		<b>57</b>
<b>5.1</b>	<b>The QSite Panel</b>	57
5.1.1	Source of Structure Input	58
5.1.2	Action Buttons	58
<b>5.2</b>	<b>The Potential Folder</b>	59
<b>5.3</b>	<b>The Constraints Folder</b>	62
<b>5.4</b>	<b>The Minimization Folder</b>	64

<b>5.5 The Optimization Folder</b> .....	66
5.5.1 Calculation Type.....	66
5.5.2 Transition State Searches.....	67
<b>5.6 The QM Settings Folder</b> .....	68
<b>5.7 Running QSite Jobs</b> .....	72
<b>Chapter 6: Running QSite from the Command Line</b> .....	<b>75</b>
<b>6.1 QSite Files</b> .....	75
<b>6.2 The qsite Command</b> .....	76
<b>6.3 Command-Line Applications and Utilities</b> .....	79
<b>Chapter 7: QSite Technical Notes</b> .....	<b>81</b>
<b>7.1 QM/MM for Protein Active Sites</b> .....	81
<b>7.2 QM/MM Transition State Modeling</b> .....	82
<b>7.3 How QSite Works</b> .....	83
<b>7.4 Parametrization Validation</b> .....	85
7.4.1 Deprotonation Energies .....	85
7.4.2 Conformational Energies .....	86
7.4.3 Other Comparisons.....	86
<b>7.5 An Illustrative Application</b> .....	87
<b>7.6 References</b> .....	88
<b>Chapter 8: Getting Help</b> .....	<b>91</b>
<b>Copyright Notices</b> .....	<b>93</b>
<b>Index</b> .....	<b>95</b>

---

# Document Conventions

In addition to the use of italics for names of documents, the font conventions that are used in this document are summarized in the table below.

*Table 3.1.*

Font	Example	Use
Sans serif	Project Table	Names of GUI features, such as panels, menus, menu items, buttons, and labels
Monospace	<code>\$SCHRODINGER/maestro</code>	File names, directory names, commands, environment variables, and screen output
Italic	<i>filename</i>	Text that the user must replace with a value
Sans serif uppercase	CTRL+H	Keyboard keys

In descriptions of command syntax, the following UNIX conventions are used: braces { } enclose a choice of required items, square brackets [ ] enclose optional items, and the pipe symbol | separates items in a list from which one item must be chosen. Lines of command syntax that wrap should be interpreted as a single command.

In this document, to *type* text means to type the required text in the specified location, and to *enter* text means to type the required text, then press the ENTER key.

References to literature sources are given in square brackets, like this: [10].



# Introduction

## 1.1 About QSite

**QSite** is a mixed mode Quantum Mechanics/Molecular Mechanics (QM/MM) program used to study geometries and energies of structures not parameterized for use with molecular mechanics, such as those that contain metals or represent transition states. QSite is uniquely equipped to perform QM/MM calculations because it combines the superior speed and power of Jaguar<sup>™</sup> with the recognized accuracy of the OPLS-AA force field. Jaguar is used for the quantum mechanical part of the calculations. The Impact<sup>™</sup> computational engine provides the molecular mechanics simulation for QSite.

QSite is run primarily from the Maestro<sup>™</sup> graphical user interface. An introduction to Maestro is provided in [Chapter 2](#) and a tutorial in using QSite from Maestro appears in [Chapter 3](#). QSite can also be run from the command line, as described in [Chapter 6](#). Utilities and scripts are also run from the command line.

**Maestro** is Schrödinger's powerful, unified, multi-platform graphical user interface (GUI). It is designed to simplify modeling tasks, such as molecule building and data analysis, and also to facilitate the setup and submission of jobs to Schrödinger's computational programs. The main Maestro features include a project-based data management facility, a scripting language for automating large or repetitive tasks, a wide range of useful display options, a comprehensive molecular builder, and surfacing and entry plotting facilities. For more detailed information about the Maestro interface than is provided in [Chapter 2](#), see the Maestro online help or the *Maestro User Manual*.

**Protein Preparation** for use in QSite can be performed for most protein and protein-ligand complex PDB structures using the Protein Preparation panel in Maestro. Python scripts and command-line utilities complete the protein preparation facility.

The Jaguar component can be run in parallel if multiple processors are available, either from the command line or from the GUI.

## 1.2 Citing QSite in Publications

The use of this product should be acknowledged in publications as:

QSite, version 4.0, Schrödinger, LLC, New York, NY, 2005.



# Introduction to Maestro

Maestro is the graphical user interface for all of Schrödinger's products: CombiGlide™, Epik™, Glide™, Impact™, Jaguar™, Liaison™, LigPrep™, MacroModel®, Phase™, Prime™, QikProp™, QSite™, and Strike™. It contains tools for building, displaying, and manipulating chemical structures; for organizing, loading, and storing these structures and associated data; and for setting up, monitoring, and visualizing the results of calculations on these structures. This chapter provides a brief introduction to Maestro and some of its capabilities. For more information on any of the topics in this chapter, see the [Maestro User Manual](#).

## 2.1 General Interface Behavior

Most Maestro panels are amodal: more than one panel can be open at a time, and a panel need not be closed for an action to be carried out. Each Maestro panel has a Close button so you can hide the panel from view.

Maestro supports the mouse functions common to many graphical user interfaces. The left button is used for choosing menu items, clicking buttons, and selecting objects by clicking or dragging. This button is also used for resizing and moving panels. The right button displays a shortcut menu. Other common mouse functions are supported, such as using the mouse in combination with the SHIFT or CTRL keys to select a range of items and select or deselect a single item without affecting other items.

In addition, the mouse buttons are used for special functions described later in this chapter. These functions assume that you have a three-button mouse. If you have a two-button mouse, ensure that it is configured for three-button mouse simulation (the middle mouse button is simulated by pressing or holding down both buttons simultaneously).

## 2.2 Starting Maestro

Before starting Maestro, you must first set the SCHRODINGER environment variable to point to the installation directory. To set this variable, enter the following command at a shell prompt:

```
cshtcsh: setenv SCHRODINGER installation-directory
bashksh: export SCHRODINGER=installation-directory
```

You might also need to set the `DISPLAY` environment variable, if it is not set automatically when you log in. To determine if you need to set this variable, enter the command:

```
echo $DISPLAY
```

If the response is a blank line, set the variable by entering the following command:

```
csh/tcsh:      setenv DISPLAY display-machine-name:0.0
```

```
bash/ksh:      export DISPLAY=display-machine-name:0.0
```

After you set the `SCHRODINGER` and `DISPLAY` environment variables, you can start Maestro using the command:

```
$SCHRODINGER/maestro options
```

If you add the `$SCHRODINGER` directory to your path, you only need to enter the command `maestro`. Options for this command are given in [Section 2.1](#) of the *Maestro User Manual*.

The directory from which you started Maestro is Maestro's current working directory, and all data files are written to and read from this directory unless otherwise specified (see [Section 2.8 on page 25](#)). You can change directories by entering the following command in the command input area (see [page 6](#)) of the main window:

```
cd directory-name
```

where *directory-name* is either a full path or a relative path.

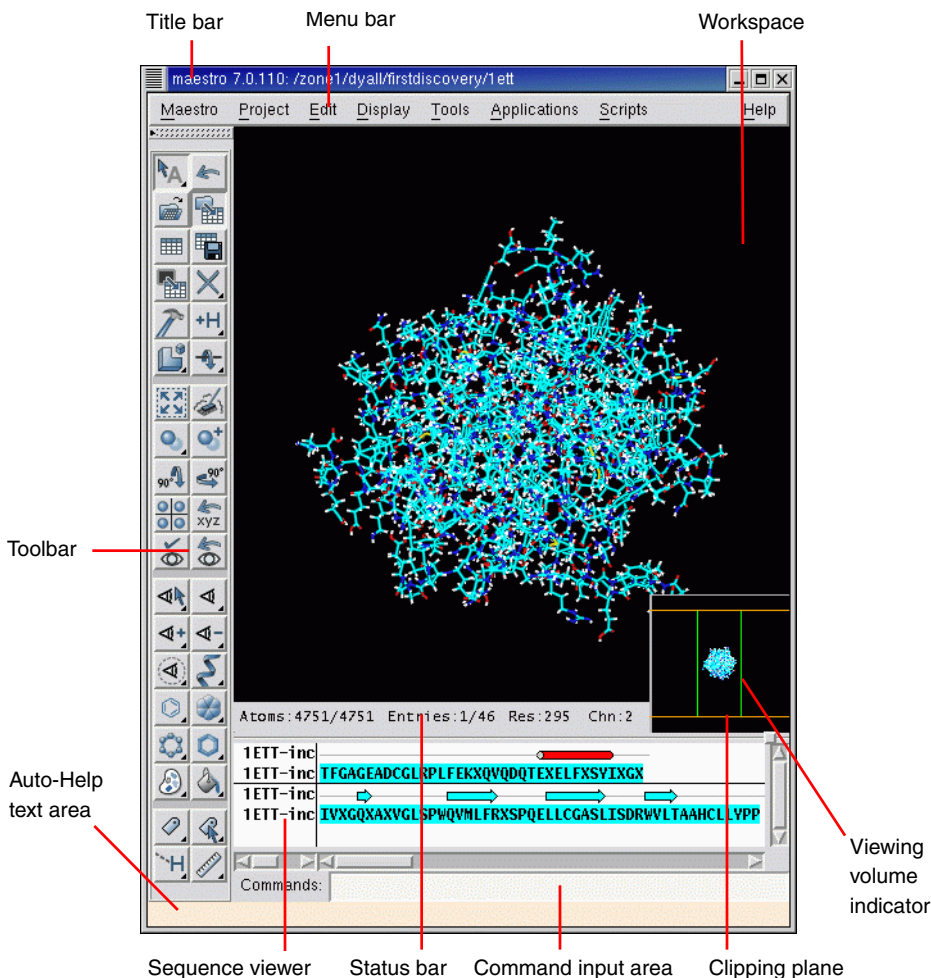
## 2.3 The Maestro Main Window

The Maestro main window is shown in [Figure 2.1 on page 5](#). The main window components are listed below.

The following components are always visible:

- **Title bar**—displays the Maestro version, the project name (if there is one) and the current working directory.
- **Auto-Help**—automatically displays context-sensitive help.
- **Menu bar**—provides access to panels.
- **Workspace**—displays molecular structures and other 3D graphical objects.

The following components can be displayed or hidden by choosing the component from the Display menu. Your choice of which main window components are displayed is persistent between Maestro sessions.



**Figure 2.1. The Maestro main window.**

- **Toolbar**—contains buttons for many common tasks and provides tools for displaying and manipulating structures, as well as organizing the Workspace.
- **Status bar**—displays information about a particular atom, or about structures in the Workspace, depending on where the pointer pauses (see [Section 2.5](#) of the *Maestro User Manual* for details):
  - **Atom**—displays the chain, residue number, element, PDB atom name, formal charge, and title or entry name (this last field is set by choosing Preferences from the Maestro menu and selecting the Feedback folder).

- **Workspace**—displays the number of atoms, entries, residues, chains, and molecules in the Workspace.
- **Clipping planes window**—displays a small, top view of the Workspace and shows the clipping planes and viewing volume indicators.
- **Sequence viewer**—shows the sequences for proteins displayed in the Workspace. See [Section 2.6](#) of the *Maestro User Manual* for details.
- **Command input area**—provides a place to enter Maestro commands.

When a distinction between components in the main window and those in other panels is needed, the term *main* is applied to the main window components (e.g., main toolbar).

You can expand the Workspace to occupy the full screen, by pressing CTRL+=. All other components and panels are hidden. To return to the previous display, press CTRL+= again.

### 2.3.1 The Menu Bar

The menus on the main menu bar provide access to panels, allow you to execute commands, and control the appearance of the Workspace. The main menus are as follows:

- **Maestro**—save or print images in the Workspace, execute system commands, save or load a panel layout, set preferences, set up Maestro command aliases, and quit Maestro.
- **Project**—open and close projects, import and export structures, make a snapshot, and annotate a project. These actions can also be performed from the Project Table panel. For more information, see [Section 2.4 on page 11](#).
- **Edit**—undo actions, build and modify structures, define command scripts and macros, and find atoms in the Workspace.
- **Display**—control the display of the contents of the Workspace, arrange panels, and display or hide main window components.
- **Tools**—group atoms; measure, align, and superimpose structures; and view and visualize data.
- **Applications**—set up, submit, and monitor jobs for Schrödinger’s computational programs. Some products have a submenu from which you can choose the task to be performed.
- **Scripts**—manage and install Python scripts that come with the distribution and scripts that you create yourself. (See [Chapter 13](#) of the *Maestro User Manual* for details.)
- **Help**—open the Help panel, the PDF documentation index, or information panels; run a demonstration; and display or hide Balloon Help (tooltips).

### 2.3.2 The Toolbar

The main toolbar contains three kinds of buttons for performing common tasks:



**Action**—Perform a simple task, like clearing the Workspace.



**Display**—Open or close a panel or open a dialog box, such as the Project Table panel.



**Menu**—Display a *button menu*. These buttons have a triangle in the lower right corner.

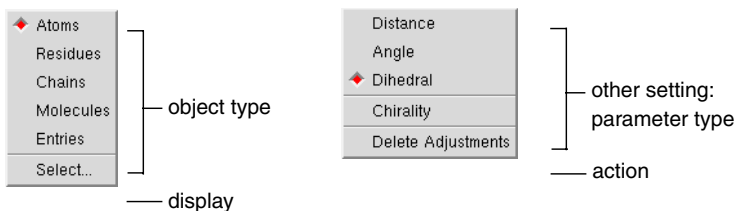
There are four types of items on button menus, and all four types can be on the same menu (see Figure 2.2):

- **Action**—Perform an action immediately.
- **Display**—Open a panel or dialog box.
- **Object types for selection**—Choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

The object type is marked on the menu with a red diamond and the button is indented to indicate the action to be performed.

- **Other setting**—Set a state, choose an attribute, or choose a parameter and click on atoms in the Workspace to display or change that parameter.

The toolbar buttons are described below. Some descriptions refer to features not described in this chapter. See the *Maestro User Manual* for a fuller description of these features.



**Figure 2.2. The Workspace selection *button menu* and the Adjust distances, angles or dihedrals *button menu*.**

### Workspace selection

- Choose an object type for selecting
- Open the Atom Selection dialog box

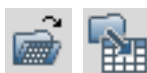


### Undo/Redo

Undo or redo the last action. Performs the same function as the Undo item on the Edit menu, and changes to an arrow pointing in the opposite direction when an Undo has been performed, indicating that its next action is Redo.

### Open a project

Open the Open Project dialog box.



### Import structures

Open the Import panel.

### Open/Close Project Table

Open the Project Table panel or close it if it is open.



### Save as

Open the Save Project As dialog box, to save the project with a new name.

### Create entry from Workspace

Open a dialog box in which you can create an entry in the current project using the contents of the Workspace.

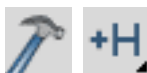


### Delete

- Choose an object type for deletion
- Delete hydrogens and waters
- Open the Atom Selection dialog box
- Delete other items associated with the structures in the Workspace
- Click to select atoms to delete
- Double-click to delete all atoms

### Open/Close Build panel

Open the Build panel or close it if it is open.



### Add hydrogens

- Choose an object type for applying a hydrogen treatment
- Open the Atom Selection dialog box
- Click to select atoms to treat
- Double-click to apply to all atoms

### Local transformation

- Choose an object type for transforming
- Click to select atoms to transform
- Open the Advanced Transformations panel



### Adjust distances, angles or dihedrals

- Choose a parameter for adjusting
- Delete adjustments

### Fit to screen

Scale the displayed structure to fit into the Workspace and reset the center of rotation.



### Clear Workspace

Clear all atoms from the Workspace.

### Set fog display state

Choose a fog state. Automatic means fog is on when there are more than 40 atoms in the Workspace, otherwise it is off.



### Enhance depth cues

Optimize fogging and other depth cues based on what is in the Workspace.

### Rotate around X axis by 90 degrees

Rotate the Workspace contents around the X axis by 90 degrees.



### Rotate around Y axis by 90 degrees

Rotate the Workspace contents around the Y axis by 90 degrees.

**Tile entries**

Arrange entries in a rectangular grid in the Workspace.

**Save view**

Save the current view of the Workspace: orientation, location, and zoom.

**Display only selected atoms**

- Choose an object type for displaying
- Click to select atoms to display
- Double-click to display all atoms

**Also display**

- Choose a predefined atom category
- Open the Atom Selection dialog box

**Display residues within N angstroms of currently displayed atoms**

- Choose a radius
- Open a dialog box to set a value

**Draw bonds in wire**

- Choose an object type for drawing bonds in wire representation
- Open the Atom Selection dialog box
- Click to select atoms for representation
- Double-click to apply to all atoms

**Draw atoms in Ball & Stick**

- Choose an object type for drawing bonds in Ball & Stick representation
- Open the Atom Selection dialog box
- Click to select atoms for representation
- Double-click to apply to all atoms

**Color all atoms by scheme**

Choose a predefined color scheme.

**Label atoms**

- Choose a predefined label type
- Delete labels

**Reset Workspace**

Reset the rotation, translation, and zoom of the Workspace to the default state.

**Restore view**

Restore the last saved view of the Workspace: orientation, location, and zoom.

**Display only**

- Choose a predefined atom category
- Open the Atom Selection dialog box

**Undisplay**

- Choose a predefined atom category
- Open the Atom Selection dialog box

**Show, hide, or color ribbons**

- Choose to show or hide ribbons
- Choose a color scheme for coloring ribbons

**Draw atoms in CPK**

- Choose an object type for drawing bonds in CPK representation
- Open the Atom Selection dialog box
- Click to select atoms for representation
- Double-click to apply to all atoms

**Draw bonds in tube**

- Choose an object type for drawing bonds in tube representation
- Open the Atom Selection dialog box
- Click to select atoms for representation
- Double-click to apply to all atoms

**Color residue by constant color**

- Choose a color for applying to residues
- Click to select residues to color
- Double-click to color all atoms

**Label picked atoms**

- Choose an object type for labeling atoms
- Open the Atom Selection dialog box
- Open the Atom Labels panel at the Composition folder
- Delete labels
- Click to select atoms to label
- Double-click to label all atoms



## Display H-bonds

- Choose bond type:  
intra—displays H-bonds within the selected molecule  
inter—displays H-bonds between the selected molecule and all other atoms.
- Delete H-bonds
- Click to select molecule



## Measure distances, angles or dihedrals

- Choose a parameter for displaying measurements
- Delete measurements
- Click to select atoms for measurement

### 2.3.3 Mouse Functions in the Workspace

The left mouse button is used for selecting objects. You can either click on a single atom or bond, or you can drag to select multiple objects. The right mouse button opens shortcut menus, which are described in [Section 2.7](#) of the *Maestro User Manual*.

The middle and right mouse buttons can be used on their own and in combination with the SHIFT and CTRL keys to perform common operations, such as rotating, translating, centering, adjusting, and zooming.

Table 2.1. Mapping of Workspace operations to mouse actions.

Mouse Button	Keyboard	Motion	Action
Left		click, drag	Select
Left	SHIFT	click, drag	Toggle the selection
Middle		drag	Rotate about X and Y axes Adjust bond, angle, or dihedral
Middle	SHIFT	drag vertically	Rotate about X axis
Middle	SHIFT	drag horizontally	Rotate about Y axis
Middle	CTRL	drag horizontally	Rotate about Z axis
Middle	SHIFT + CTRL	drag horizontally	Zoom
Right		click	Spot-center on selection
Right		click and hold	Display shortcut menu
Right		drag	Translate in the X-Y plane
Right	SHIFT	drag vertically	Translate along the X axis
Right	SHIFT	drag horizontally	Translate along the Y axis
Right	CTRL	drag horizontally	Translate along the Z axis
Middle & Right		drag horizontally	Zoom

### 2.3.4 Shortcut Key Combinations

Some frequently used operations have been assigned shortcut key combinations. The shortcuts available in the main window are described in [Table 2.2](#).

Table 2.2. *Shortcut keys in the Maestro main window.*

Keys	Action	Equivalent Menu Choices
CTRL+B	Open Build panel	Edit > Build
CTRL+C	Create entry	Project > Create Entry From Workspace
CTRL+E	Open Command Script Editor panel	Edit > Command Script Editor
CTRL+F	Open Find Atoms panel	Edit > Find
CTRL+H	Open Help panel	Help > Help
CTRL+I	Open Import panel	Project > Import Structures
CTRL+M	Open Measurements panel	Tools > Measurements
CTRL+N	Create new project	Project > New
CTRL+O	Open project	Project > Open
CTRL+P	Print	Maestro > Print
CTRL+Q	Quit	Maestro > Quit
CTRL+S	Open Sets panel	Tools > Sets
CTRL+T	Open Project Table panel	Project > Show Table
CTRL+W	Close project	Project > Close
CTRL+Z	Undo/Redo last command	Edit > Undo/Redo
CTRL+=	Enter and exit full screen mode (Workspace occupies full screen)	None

## 2.4 Maestro Projects

All the work you do in Maestro is done within a *project*. A project consists of a set of *entries*, each of which contains one or more chemical structures and their associated data. In any Maestro session, there can be only one Maestro project open. If you do not specify a project when you start Maestro, a *scratch* project is created. You can work in a scratch project without saving it, but you must save it in order to use it in future sessions. When you save or close a project, all the view transformations (rotation, translation, and zoom) are saved with it. When you close a project, a new scratch project is automatically created.

Likewise, if there is no entry displayed in the Workspace, Maestro creates a *scratch* entry. Structures that you build in the Workspace constitute a scratch entry until you save the structures as project entries. The scratch entry is not saved with the project unless you explicitly add it to the project. However, you can use a scratch entry as input for some calculations.

To add a scratch entry to a project, do one of the following:

- Click the Create entry from Workspace button:



- Choose Create Entry from Workspace from the Project menu.
- Press CTRL+C.

In the dialog box, enter a name and a title for the entry. The entry name is used internally to identify the entry and can be modified by Maestro. The title can be set or changed by the user, but is not otherwise modified by Maestro.

Once an entry has been incorporated into the project, its structures and their data are represented by a row in the Project Table. Each row contains the row number, an icon indicating whether the entry is displayed in the Workspace (the In column), the entry title, a button to open the Surfaces panel if the entry has surfaces, the entry name, and any entry properties. The row number is not a property of the entry.

Entries can be collected into groups, and the members of the group can be displayed or hidden. Most additions of multiple entries to the Project Table are done as entry groups.

You can use entries as input for all of the computational programs—Glide, Impact, Jaguar, Liaison, LigPrep, MacroModel, Phase, Prime, QikProp, QSite, and Strike. You can select entries as input for the ePlayer, which displays the selected structures in sequence. You can also duplicate, combine, rename, and sort entries; create properties; import structures as entries; and export structures and properties from entries in various formats.

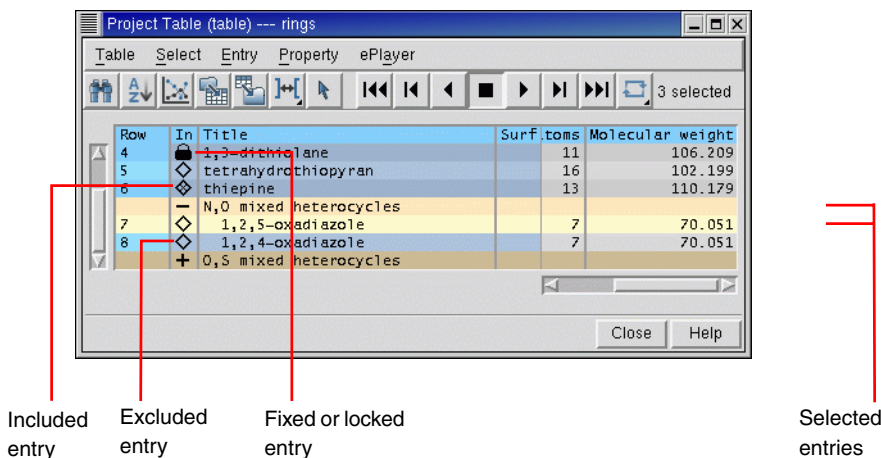
To open the Project Table panel, do one of the following:

- Click the Open/Close Project Table button on the toolbar



- Choose Show Table from the Project menu
- Press CTRL+T.

The Project Table panel contains a menu bar, a toolbar, and the table itself.



**Figure 2.3.** The Project Table *panel*.

### 2.4.1 The Project Table Toolbar

The Project Table toolbar contains two groups of buttons and a status display. The first set of buttons opens various panels that allow you to perform functions on the entries in the Project Table. The second set of buttons controls the ePlayer, which “plays through” the selected structures: each structure is displayed in the Workspace in sequence, at a given time interval. See [Section 2.3.2 on page 7](#) for a description of the types of toolbar buttons. The buttons are described below.



#### Find

Open the Find panel for locating alphanumeric text in any column of the Project Table, except for the row number.



#### Sort

Open the Sort panel for sorting entries by up to three properties.



#### Plot

Open the Plot panel for plotting entry properties.



#### Import Structure

Open the Import panel for importing structures into the project.



#### Export Structure

Open the Export panel for exporting structures to a file.



**Columns**

Choose an option for adjusting the column widths.



**Select only**

Open the Entry Selection dialog box for selecting entries based on criteria for entry properties.



**Go to start**

Display the first selected structure.



**Previous**

Display the previous structure in the list of selected structures.



**Play backward**

Display the selected structures in sequence, moving toward the first.



**Stop**

Stop the ePlayer.



**Play forward**

Display the selected structures in sequence, moving toward the last.



**Next**

Display the next structure in the list of selected structures.



**Go to end**

Display the last selected structure.



**Loop**

Choose an option for repeating the display of the structures. **Single Direction** displays structures in a single direction, then repeats. **Oscillate** reverses direction each time the beginning or end of the list is reached.

The status display, to the right of the toolbar buttons, shows the number of selected entries. When you pause the cursor over the status display, the Balloon Help shows the total number of entries, the number shown in the table, the number selected, and the number included in the Workspace.

## 2.4.2 The Project Table Menus

- **Table**—find text, sort entries, plot properties, import and export structures, and configure the Project Table.
- **Select**—select all entries, none, invert your selection, or select classes of entries using the Entry Selection dialog box and the Filter panel.


- **Entry**—include or exclude entries from the Workspace, display or hide entries in the Project Table, and perform various operations on the selected entries.
- **Property**—display and manipulate entry properties in the Project Table.
- **ePlayer**—view entries in succession, stop, reverse, and set the ePlayer options.

### 2.4.3 Selecting Entries

Many operations in Maestro are performed on the entries selected in the Project Table. The Project Table functions much like any other table: select rows by clicking, shift-clicking, and control-clicking. However, because clicking in an editable cell of a selected row enters edit mode, you should click in the Row column to select entries. See [Section 2.4.5 on page 16](#) for more information on mouse actions in the Project Table. There are shortcuts for selecting classes of entries on the **Select** menu.

In addition to selecting entries manually, you can select entries that meet a combination of conditions on their properties. Such combinations of conditions are called *filters*. Filters are Entry Selection Language (ESL) expressions and are evaluated at the time they are applied. For example, if you want to set up a Glide job that uses ligands with a low molecular weight (say, less than 300) and that has certain QikProp properties, you can set up a filter and use it to select entries for the job. If you save the filter, you can use it again on a different set of ligands that meet the same selection criteria.

#### To create a filter:

1. Do one of the following:
    - Choose **Only**, **Add**, or **Deselect** from the **Select** menu.
    - Click the **Entry selection** button on the toolbar.
- 
2. In the **Properties** folder, select a property from the property list, then select a condition.
  3. Combine this selection with the current filter by clicking **Add**, **Subtract**, or **Intersect**. These buttons perform the Boolean operations **OR**, **AND NOT**, and **AND** on the corresponding ESL expressions.
  4. To save the filter for future use click **Create Filter**, enter a name, and click **OK**.
  5. Click **OK** to apply the filter immediately.

### 2.4.4 Including Entries in the Workspace

In addition to selecting entries, you can also use the Project Table to control which entries are displayed in the Workspace. An entry that is displayed in the Workspace is *included* in the Workspace; likewise, an entry that is not displayed is *excluded*. Included entries are marked by an X in the diamond in the In column; excluded entries are marked by an empty diamond. Entry inclusion is completely independent of entry selection.

To include or exclude entries, click, shift-click, or control-click in the In column of the entries, or select entries and choose Include or Exclude from the Entry menu. Inclusion with the mouse works just like selection: when you include an entry by clicking, all other entries are excluded.

It is sometimes useful to keep one entry in the Workspace and include others one by one: for example, a receptor and a set of ligands. You can fix the receptor in the Workspace by selecting it in the Project Table and choosing Fix from the Entry menu or by pressing CTRL+F. A padlock icon replaces the diamond in the In column to denote a *fixed* entry. To remove a fixed entry from the Workspace, you must exclude it explicitly (CTRL+X). It is not affected by the inclusion or exclusion of other entries. Fixing an entry affects only its inclusion; you can still rotate, translate, or modify the structure.

### 2.4.5 Mouse Functions in the Project Table

The Project Table supports the standard use of shift-click and control-click to select objects. This behavior applies to the selection of entries and the inclusion of entries in the Workspace. You can also drag to resize rows and columns and to move rows.

You can drag a set of non-contiguous entries to reposition them in the Project Table. When you release the mouse button, the entries are placed after the first unselected entry that precedes the entry on which the cursor is resting. For example, if you select entries 2, 4, and 6, and release the mouse button on entry 3, these three entries are placed after entry 1, because entry 1 is the first unselected entry that precedes entry 3. To move entries to the top of the table, drag them above the top of the table; to move entries to the end of the table, drag them below the end of the table.

A summary of mouse functions in the Project Table is provided in [Table 2.3](#).

Table 2.3. Mouse operations in the Project Table.

Task	Mouse Operation
Change a Boolean property value	Click repeatedly in a cell to cycle through the possible values (On, Off, Clear)
Display the Entry menu for an entry	Right-click anywhere in the entry. If the entry is not selected, it becomes the selected entry. If the entry is selected, the action is applied to all selected entries.
Display a version of the Property menu for a property	Right-click in the column header
Edit the text or the value in a table cell	Click in the cell and edit the text or value
Include an entry in the Workspace, exclude all others	Click the In column of the entry
Move selected entries	Drag the entries
Paste text into a table cell	Middle-click
Resize rows or columns	Drag the boundary with the middle mouse button
Select an entry, deselect all others	For an unselected entry, click anywhere in the row except the In column; for a selected entry, click the row number.
Select or include multiple entries	Click the first entry then shift-click the last entry
Toggle the selection or inclusion state	Control-click the entry or the In column

## 2.4.6 Project Table Shortcut Keys

Some frequently used project operations have been assigned shortcut key combinations. The shortcuts, their functions, and their menu equivalents are listed in [Table 2.4](#).

Table 2.4. Shortcut keys in the Project Table.

Keys	Action	Equivalent Menu Choices
CTRL+A	Select all entries	Select > All
CTRL+F	Fix entry in Workspace	Entry > Fix
CTRL+I	Open Import panel	Table > Import Structures
CTRL+N	Include only selected entries	Entry > Include Only
CTRL+U	Deselect all entries	Select > None
CTRL+X	Exclude selected entries	Entry > Exclude
CTRL+Z	Undo/Redo last command	Edit > Undo/Redo in main window

## 2.5 Building a Structure

After you start Maestro, the first task is usually to create or import a structure. You can open existing Maestro projects or import structures from other sources to obtain a structure, or you can build your own. To open the Build panel, do one of the following:

- Click the Open/Close Build panel button in the toolbar:



- Choose Build from the Edit menu.
- Press CTRL+B.

The Build panel allows you to create structures by drawing or placing atoms or fragments in the Workspace and connecting them into a larger structure, to adjust atom positions and bond orders, and to change atom properties. This panel contains a toolbar and three folders.

### 2.5.1 Placing and Connecting Fragments

The Build panel provides several tools for creating structures in the Workspace. You can place and connect fragments, or you can draw a structure freehand.

#### To place a fragment in the Workspace:

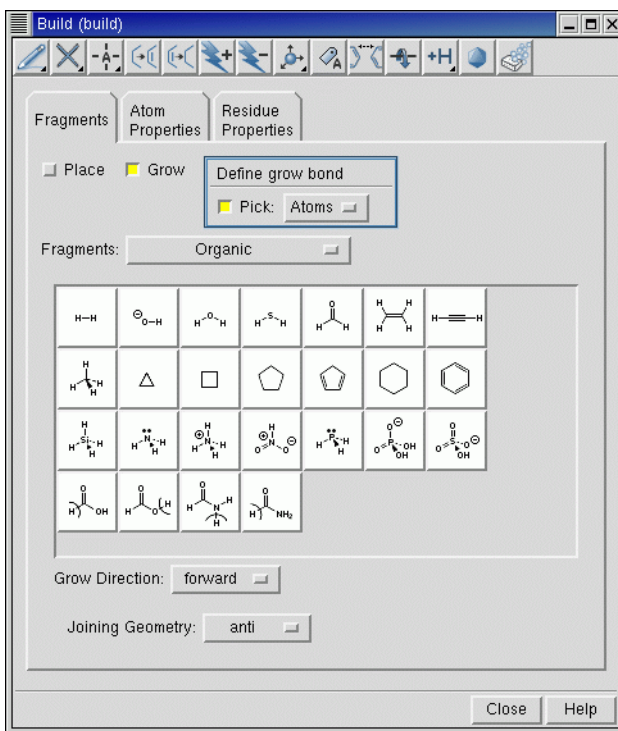
1. Select Place.
2. Choose a fragment library from the Fragments menu.
3. Click a fragment.
4. Click in the Workspace where you want the fragment to be placed.

#### To connect fragments in the Workspace, do one of the following:

- Place another fragment and connect them using the Connect & Fuse panel, which you open from the Edit menu on the main menu bar or with the Display Connect & Fuse panel on the Build toolbar.



- Replace one or more atoms in the existing fragment with another fragment by selecting a fragment and clicking in the Workspace on the main atom to be replaced.
- Grow another fragment by selecting Grow in the Build panel and clicking the fragment you want to add in the Fragments folder.



**Figure 2.4. The Build panel.**

Grow mode uses predefined rules to connect a fragment to the *grow bond*. The grow bond is marked by a green arrow. The new fragment replaces the atom at the head of the arrow on the grow bond and all atoms attached to it. To change the grow bond, choose Bonds from the Pick option menu in the Build panel and click on the desired grow bond in the Workspace. The arrow points to the atom nearest to where you clicked.

#### To draw a structure freehand:

1. Choose an element from the Draw button menu on the Build panel toolbar:



2. Click in the Workspace to place an atom of that element.
3. Click again to place another atom and connect it to the previous atom.
4. Continue this process until you have drawn the structure.
5. Click the active atom again to finish drawing.

### 2.5.2 Adjusting Properties

In the Atom Properties folder, you can change the properties of the atoms in the Workspace. For each item on the Property option menu—Element, Atom Type (MacroModel), Partial Charge, PDB Atom Name, Grow Name, and Atom Name—there is a set of tools you can use to change the atom properties. For example, the Element tools consist of a periodic table from which you can choose an element and select an atom to change it to an atom of the selected element.

Similarly, the Residue Properties folder provides tools for changing the properties of residues: the Residue Number, the Residue Name, and the Chain Name.

To adjust bond lengths, bond angles, dihedral angles, and chiralities during or after building a structure, use the Adjust distances, angles or dihedrals button on the main toolbar:



You can also open the Adjust panel from this button menu, from the Display Adjust panel button on the Build panel toolbar (which has the same appearance as the above button) or from the Edit menu in the main window.

### 2.5.3 The Build Panel Toolbar

The toolbar of the Build panel provides quick access to tools for drawing and modifying structures and labeling atoms. See [Section 2.3.2 on page 7](#) for a description of the types of toolbar buttons. The toolbar buttons and their use are described below.



#### Free-hand drawing

Choose an element for drawing structures freehand in the Workspace (default C). Each click in the Workspace places an atom and connects it to the previous atom.



#### Delete

Choose an object for deleting. Same as the [Delete](#) button on the main toolbar, see [page 8](#).



#### Set element

Choose an element for changing atoms in the Workspace (default C). Click an atom to change it to the selected element.



#### Increment bond order

Select a bond to increase its bond order by one, to a maximum of 3.



#### Decrement bond order

Select a bond to decrease its bond order by one, to a minimum of 0.

**Increment formal charge**

Select an atom to increase its formal charge by one.

**Decrement formal charge**

Select an atom to decrease its formal charge by one.

**Move**

Choose a direction for moving atoms, then click the atom to be moved. Moves in the XY plane are made by clicking the new location. Moves in the Z direction are made in 0.5 Å increments.

**Label**

Apply heteroatom labels as you build a structure. The label consists of the element name and formal charge, and is applied to atoms other than C and H.

**Display Connect & Fuse panel**

Open the Connect & Fuse panel so you can connect structures (create bonds between structures) or fuse structures (replace atoms of one structure with those of another).

**Display Adjust panel**

Open the Adjust panel so you can change bond lengths, bond angles, dihedral angles, or atom chiralities.

**Add hydrogens**

Choose an atom type for applying the current hydrogen treatment. Same as the [Add hydrogens](#) button on the main toolbar, see [page 8](#).

**Geometry Symmetrizer**

Open the Geometry Symmetrizer panel for symmetrizing the geometry of the structure in the Workspace.

**Geometry Cleanup**

Clean up the geometry of the structure in the Workspace.

## 2.6 Selecting Atoms

Maestro has a powerful set of tools for selecting atoms in a structure: toolbar buttons, picking tools in panels, and the Atom Selection dialog box. These tools allow you to select atoms in two ways:

- Select atoms first and apply an action to them
- Choose an action first and then select atoms for that action

### 2.6.1 Toolbar Buttons

The small triangle in the lower right corner of a toolbar button indicates that the button contains a menu. Many of these buttons allow you to choose an object type for selecting: choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

For example, to select atoms with the Workspace selection toolbar button:

1. Choose Residues from the Workspace selection button menu:



The button changes to:



2. Click on an atom in a residue in the Workspace to select all the atoms in that residue.

## 2.6.2 Picking Tools

The picking tools are embedded in each panel in which you need to select atoms to apply an operation. The picking tools in a panel can include one or more of the following:

- Pick option menu—Allows you to choose an object type. Depending on the operation to be performed, you can choose Atoms, Bonds, Residues, Chains, Molecules, or Entries, then click on an atom in the Workspace to perform the action on all the atoms in that structural unit.

The Pick option menu varies from panel to panel, because not all object types are appropriate for a given operation. For example, some panels have only Atoms and Bonds in the Pick option menu.

- All button—Performs the action on all atoms in the Workspace.
- Selection button—Performs the action on any atoms already selected in the Workspace.
- Previous button—Performs the action on the most recent atom selection defined in the Atom Selection dialog box.
- Select button—Opens the Atom Selection dialog box.
- ASL text box—Allows you to type in an ASL expression for selecting atoms.

ASL stands for Atom Specification Language, and is described in detail in the [Maestro Command Reference Manual](#).

- Clear button—Clears the current selection



- Show markers option—Marks the selected atoms in the Workspace.

For example, to label atoms with the Label Atoms panel:

1. Choose Atom Labels from the Display menu.
2. In the Composition folder, select Element and Atom Number.
3. In the picking tools section at the top of the panel, you could do one of the following:
  - Click Selection to apply labels to the atoms already selected in the Workspace (from the previous example).
  - Choose Residues from the Pick option menu and click on an atom in a different residue to label all the atoms in that residue.

### 2.6.3 The Atom Selection Dialog Box

If you wish to select atoms based on more complex criteria, you can use the Atom Selection dialog box. To open this dialog box, choose Select from a button menu or click the Select button in a panel. See [Section 5.3](#) of the *Maestro User Manual* for detailed instructions on how to use the Atom Selection dialog box.

## 2.7 Scripting in Maestro

Although you can perform nearly all Maestro-supported operations through menus and panels, you can also perform operations using Maestro commands, or compilations of these commands, called *scripts*. Scripts can be used to automate lengthy procedures or repetitive tasks and can be created in several ways. These are summarized below.

### 2.7.1 Python Scripts

Python is a full-featured scripting language that has been embedded in Maestro to extend its scripting facilities. The Python capabilities within Maestro include access to Maestro functionality for dealing with chemical structures, projects, and Maestro files.

The two main Python commands used in Maestro are:

- `pythonrun`—executes a Python module. (You can also use the alias `pyrun`.) The syntax is:

```
pythonrun module.function
```

- `pythonimport`—rereads a Python file so that the next time you use the `pythonrun` command, it uses the updated version of the module. (You can also use the alias `pyimp`.)

From the Maestro Scripts menu you can install, manage, and run Python scripts. For more information on the Scripts menu, see [Section 13.1](#) of the *Maestro User Manual*.

For more information on using Python with Maestro, see *Maestro Scripting with Python*.

## 2.7.2 Command Scripts

All Maestro commands are logged and displayed in the Command Script Editor panel. This means you can create a command script by performing the operations with the GUI controls, copying the logged commands from the Command History list into the Script text area of the panel, then saving the list of copied commands as a script.

### To run an existing command script:

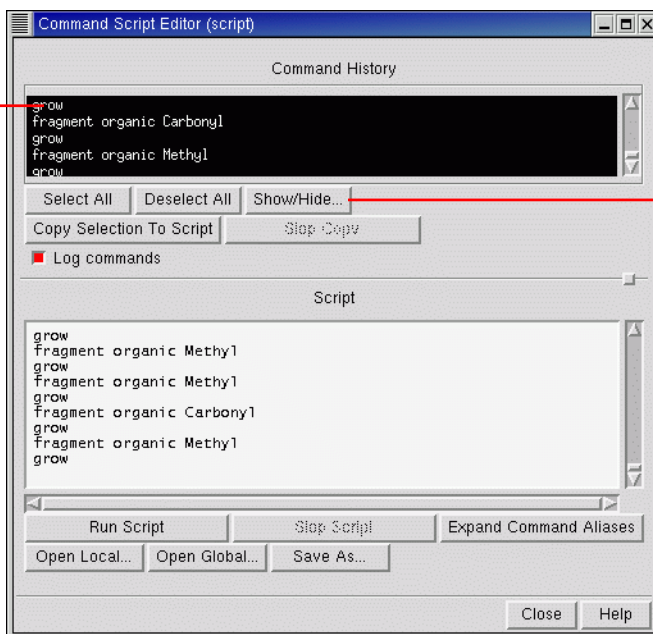
1. Open the Command Script Editor panel from the Edit menu in the main window.
2. Click Open Local and navigate to the directory containing the desired script.
3. Select a script in the Files list and click Open.

The script is loaded into the Script window of the Command Script Editor panel.

4. Click Run Script.

Command scripts cannot be used for Prime operations.

The *Command History* window displays a log of all commands issued internally within Maestro when you interact with a panel, menu, or structure



Opens the *Show/Hide Command* panel, used to determine which commands are logged in the *Command History* list

**Figure 2.5.** The Command Script Editor panel.

### 2.7.3 Macros

There are two kinds of macros you can create: named macros and macros assigned to function keys F1 through F12.

**To create and run a named macro:**

1. Open the Macros panel from the Edit menu in the main window.
2. Click New, enter a name for the macro, and click OK.
3. In the Definition text box, type the commands for the macro.
4. Click Update to update the macro definition.
5. To run the macro, enter the following in the command input area in the main window:

```
macrorun macro-name
```

If the command input area is not visible, choose Command Input Area from the Display menu.

**To create and run a function key macro:**

1. Open the Function Key Macros panel from the Edit menu in the main window.
2. From the Macro Key option, select a function key (F1 through F12) to which to assign the macro.
3. In the text box, type the commands for the macro.
4. Click Run to test the macro or click Save to save it.
5. To run the macro from the main window, press the assigned function key.

For more information on macros, see [Section 13.5](#) of the *Maestro User Manual*.

## 2.8 Specifying a Maestro Working Directory

When you use Maestro to launch QSite jobs, Maestro writes job output to the directory specified in the Directory folder of the Preferences panel. By default, this directory (the file I/O directory) is the directory from which you started Maestro.

**To change the Maestro working directory:**

1. Open the Preferences panel from the Maestro menu.
2. Click the Directory tab.
3. Select the directory you want to use for reading and writing files.

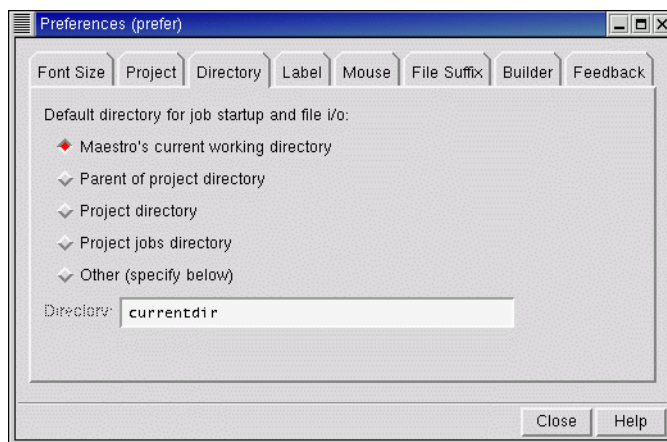


Figure 2.6. T

You can also set other preferences in the Preferences panel. See [Section 12.2](#) of the *Maestro User Manual* for details.

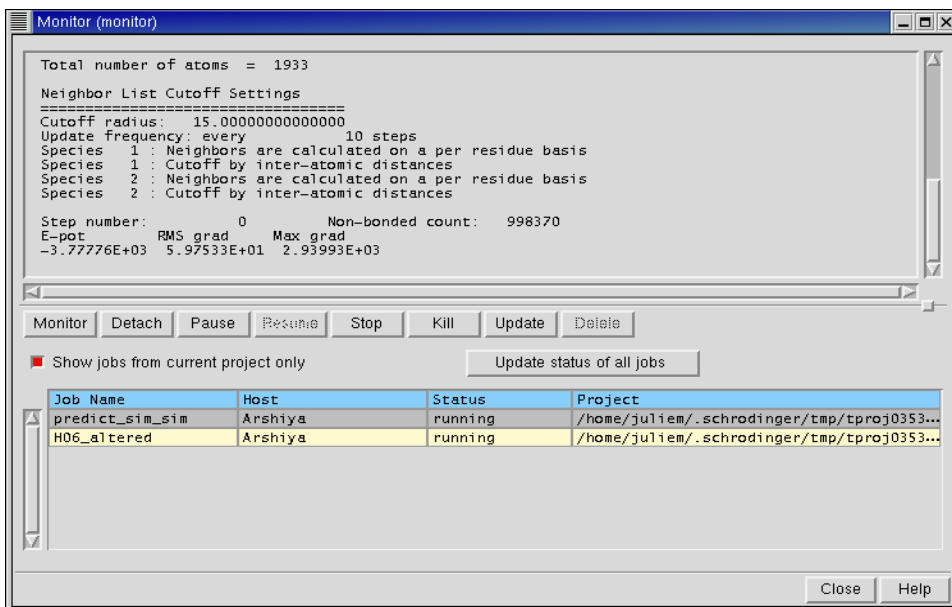
## 2.9 Undoing an Operation

To undo a single operation, click the Undo button in the toolbar, choose Undo from the Edit menu, or press CTRL+Z. The word Undo in the menu is followed by text that describes the operation to undo. Not all operations can be undone: for example, global rotations and translations are not undoable operations. For such operations you can use the Save view and Restore view buttons in the toolbar, which save and restore a molecular orientation.

## 2.10 Running and Monitoring Jobs

Maestro has panels for each product for preparing and submitting jobs. To use these panels, choose the appropriate product and task from the Applications menu and its submenus. Set the appropriate options in the panel, then click Start to open the Start dialog box and set options for running the job. For a complete description of the Start dialog box associated with your computational program, see your product's User Manual. When you have finished setting the options, click Start to launch the job and open the Monitor panel.

The Monitor panel is the control panel for monitoring the progress of jobs and for pausing, resuming, or killing jobs. All jobs that belong to your user ID can be displayed in the Monitor panel, whether or not they were started from Maestro. Subjobs are indented under their parent in the job list. The text pane shows various output information from the monitored job, such as the contents of the log file. The Monitor panel opens automatically when you start a job. If it is



**Figure 2.7. The Monitor panel.**

not open, you can open it by choosing Monitor from the Applications menu in the Maestro main window.

While jobs are running, the Detach, Pause, Resume, Stop, Kill, and Update buttons are active. When there are no jobs currently running, only the Monitor and Delete buttons are active. These buttons act on the selected job. By default, only jobs started from the current project are shown. To show other jobs, deselect Show jobs from current project only.

When a monitored job ends, the results are incorporated into the project according to the settings used to launch the job. If a job that is not currently being monitored ends, you can select it in the Monitor panel and click Monitor to incorporate the results. Monitored jobs are incorporated only if they are part of the current project. You can monitor jobs that are not part of the current project, but their results are not incorporated. To add their results to a project, you must open the project and import the results.

Further information on job control, including configuring your site, monitoring jobs, running jobs, and job incorporation, can be found in the [Job Control Guide](#) and the [Installation Guide](#).

## 2.11 Getting Help

Maestro comes with automatic, context-sensitive help (Auto-Help), Balloon Help (tooltips), an online help facility, and a user manual. To get help, follow the steps below:

- Check the Auto-Help text box at the bottom of the main window. If help is available for the task you are performing, it is automatically displayed there. It describes what actions are needed to perform the task.
- If your question concerns a GUI element, such as a button or option, there may be Balloon Help for the item. Pause the cursor over the element. If the Balloon Help does not appear, check that Show Balloon Help is selected in the Help menu of the main window. If there is Balloon Help for the element, it appears within a few seconds.
- If you do not find the help you need using either of the steps above, click the Help button in the lower right corner of the appropriate panel. The Help panel is displayed with a relevant help topic.
- For help with a concept or action not associated with a panel, open the Help panel from the Help menu or press CTRL+H.

If you do not find the information you need in the Maestro help system, check the following sources:

- The *Maestro User Manual*
- The Frequently Asked Questions page, found at <http://www.schrodinger.com/Support/faq.html>

You can also contact Schrödinger by e-mail or phone for help:

- E-mail: [help@schrodinger.com](mailto:help@schrodinger.com)
- Phone: (503) 299-1150

## 2.12 Ending a Maestro Session

To end a Maestro session, choose Quit from the Maestro menu. To save a log file with a record of all operations performed in the current session, click Quit, save log file in the Quit panel. This information can be useful to Schrödinger support staff when responding to any problem you report.

# QSite Tutorial

This chapter contains a tutorial designed to help you quickly become familiar with the functionality of QSite using the Maestro interface. In this chapter, you will perform a QSite geometry minimization on a protein-ligand complex. Density functional theory (DFT) will be used to treat the QM region, which will consist of the ligand only. This is a straightforward example of using QSite to model a stationary state. QSite can also be useful in modeling enzymatic systems involving transition states or metal atoms, which can be poorly treated by empirical force fields such as OPLS-AA.

To do these exercises, you must have access to an installed version of Maestro 7.5 and QSite 4.0. For installation instructions, see the [Installation Guide](#).

## 3.1 Preparing a Working Directory

Files needed for this tutorial are included with the Impact distribution. The `$SCHRODINGER/impact-vversion/tutorial/structures` directory contains the structure files needed to begin this tutorial. Some exercises in this tutorial produce files that are needed in subsequent exercises. To allow you to begin at any exercise you choose, the `$SCHRODINGER/impact-vversion/tutorial/qsite` directory contains copies of the input and output files that will be generated as you perform the tutorial. These files can be used as shortcuts if you want to work on a later exercise in the tutorial without performing all the previous exercises.

Before you start Maestro, you must set the `SCHRODINGER` environment variable, create a local working directory, then link to the structures in the Impact distribution.

1. Set the `SCHRODINGER` environment variable to the product installation directory:

**csh/tcsh:**            `setenv SCHRODINGER installation_path`

**sh/bash/ksh:**        `export SCHRODINGER=installation_path`

2. Change to a directory in which you have write permission.
3. Create a new working directory by entering the command:

```
mkdir qsite-workdir
```

4. In your working directory, create a soft link to the structures directory by entering:

```
ln -s $SCHRODINGER/impact-vversion/tutorial/structures .
```

## 3.2 Starting Maestro

You do not need to start Maestro until you begin the exercises. If you have not started Maestro before, this section contains instructions.

1. Change to your *qsite-workdir* directory.
2. Enter the command:

```
$SCHRODINGER/maestro &
```

The Maestro main window is displayed.

## 3.3 Importing the Complex

Use the following steps to import the protein-ligand complex in the file `1tpb.mae`. Both the protein and the ligand have been prepared for use in QSite as described in [Chapter 4](#). Bond orders are correct and hydrogens have been added.

1. On the toolbar, click the Import structures button.



The Import panel is displayed.

2. Choose Maestro from the Format menu.
3. Navigate to *qsite-workdir/structures* and select the file `1tpb.mae`.
4. Ensure that Replace workspace is selected.
5. Click Import.

The 1TPB\_2 receptor-ligand complex is included in the Workspace. The ligand is colored turquoise.

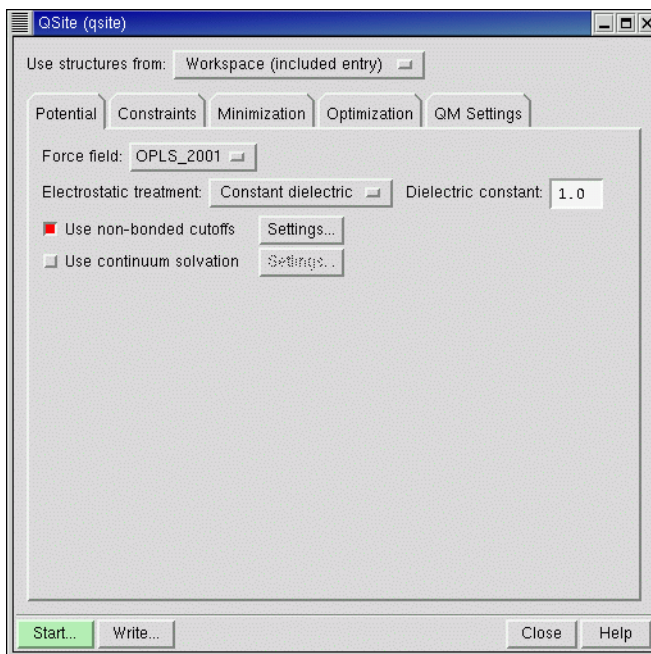
6. Click Close to close the Import panel.
7. Rotate the structure (middle mouse) so that you can see the ligand clearly.

## 3.4 Selecting the QSite Job Type

Follow the instructions below to open the QSite panel and set the job type.

1. In the main window, choose QSite from the Applications menu.

The QSite panel is displayed.



**Figure 3.1. The Potential folder of the QSite panel.**

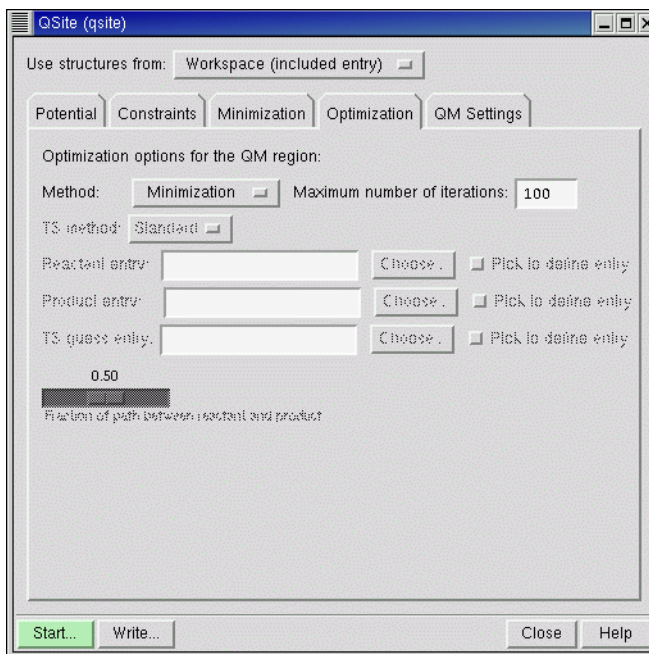
2. In the Potential folder, ensure that OPLS\_2001 is selected in the Force field option menu.
3. In the Optimization folder, choose Minimization from the Method option menu.

### 3.5 Defining a QM Region

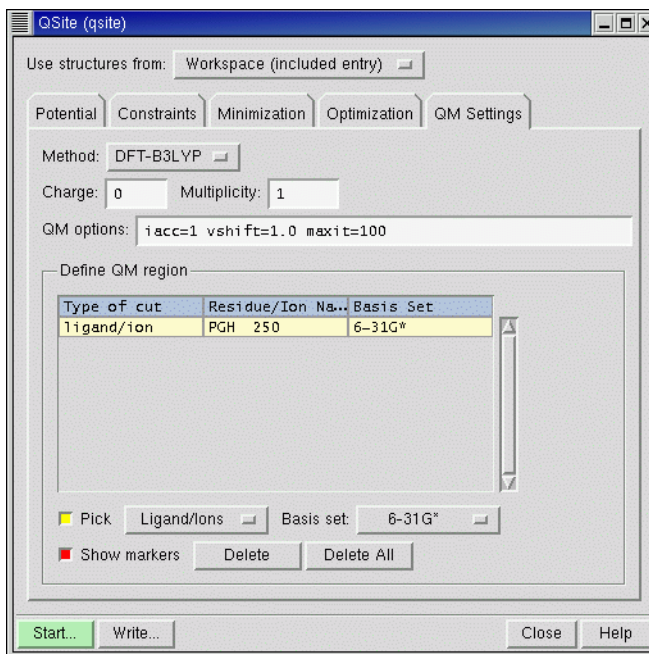
You can select an isolated ligand molecule, or a lone ion or metallic cofactor, for the QM region by simple picking. To select entire residues from protein chains, you must make back-bone cuts. It is often useful to make side chain cuts, adding only the side chain rather than the entire residue to the QM region. The following exercise demonstrates QM region definition by ligand picking.

1. In the QM Settings folder, select the Pick option.
2. Choose Ligand/Ions from the Pick option menu.
3. Zoom in on the ligand in the Workspace (by dragging horizontally with the middle and right mouse buttons).

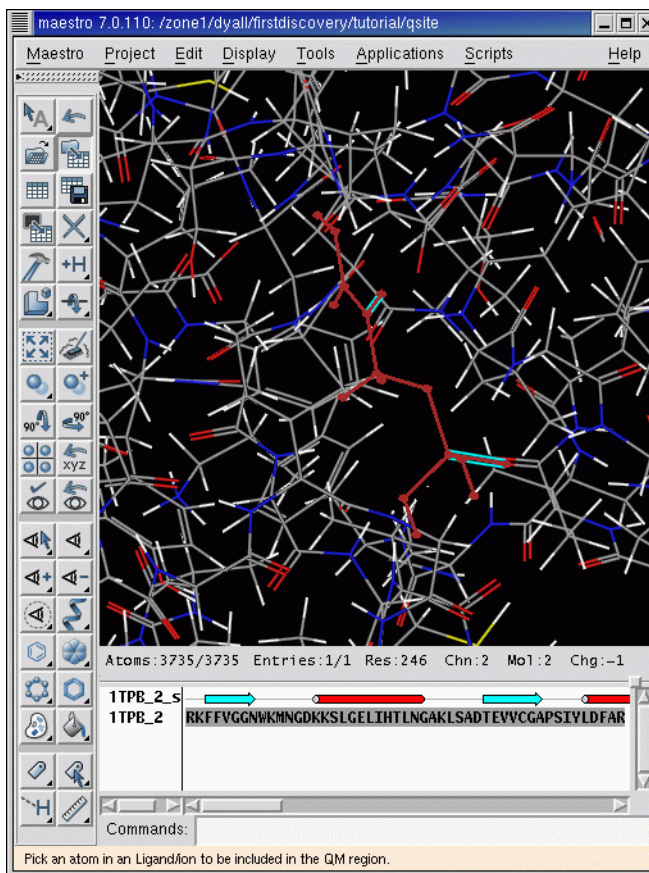
To center the display on the ligand, right-click on a ligand atom. When the cursor is over a ligand atom, the residue name appears with the atom identification in the status bar. For example, 2 : PGH 250 (C) C1.



**Figure 3.2. The Optimization *folder* of the QSite *panel*.**



**Figure 3.3. The QM Settings *folder* with the ligand picked.**



**Figure 3.4. The ligand molecule marked after selection.**

4. Pick an atom in the ligand.

Ball and stick markers in sienna brown are superimposed on the ligand molecule.

In the Define QM region table, the Type of cut, the Name of the residue or ion, and the Basis Set are listed.

QM region size is the most influential factor in QSite calculation speed. It is therefore not advantageous to work with smaller model proteins.

## 3.6 Running the QSite Job

1. In the lower left corner of the QSite panel, click Start.

The QSite - Start dialog box is displayed.

2. Change the job name to `qsite_1tpb`.
3. Click Start.

The Monitor panel opens and displays the job log as it runs.

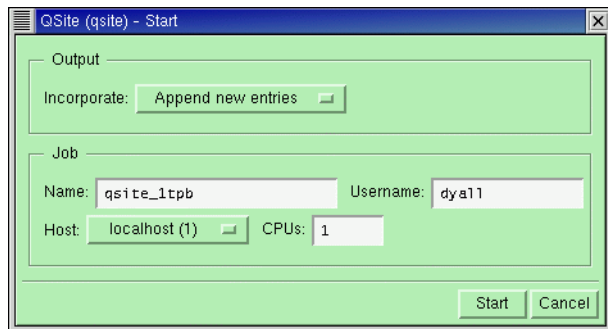
This job takes several hours of CPU time on a 1 GHz Pentium IV processor. If you wish to examine the results without waiting for the job to finish, you can import the input and output structures from the directory `$SCHRODINGER/impact-vversion/tutorial/qsite` and proceed to the next section.

The following files appear in your current Maestro working directory (*qsite-workdir*) before the job starts:

<code>qsite_1tpb.inp</code>	QSite Impact input file
<code>qsite_1tpb.mae</code>	Maestro structure file
<code>qsite_1tpb.jaguar.in</code>	Jaguar keyword file

When the job is complete, these files are written:

<code>qsite_1tpb.out</code>	QSite Impact output file
<code>qsite_1tpb.jaguar.log</code>	Log file as displayed in the Monitor panel
<code>qsite_1tpb.jaguar.01.in</code>	Jaguar restart file
<code>qsite_1tpb_out.mae</code>	Maestro structure file with optimized structure
<code>qsite_1tpb.jaguar.out</code>	Detailed output file with grid points and orbital energies



**Figure 3.5.** The QSite - Start *dialog box*.

## 3.7 Examining Results

### 3.7.1 Comparing Input and Output Structures

1. Click the Open/Close project table toolbar button



The Project Table panel opens. The output structure has been appended to the project as an entry, with properties QM/MM Energy, QM Basis, QM Method, and Job Name. By default, the output structure is included in the Workspace.

2. Control-click the check box in the In column for the original structure.

The original structure is included in the Workspace as well.

3. Choose Molecule Number (Carbons) from the Color all atoms by scheme button menu



Each of the four molecules in the Workspace (two receptors and two ligands) is now distinctly colored.

### 3.7.2 Comparing Ligand-Receptor Interactions

In this exercise you will compare the hydrogen bonding interactions between ligand and receptor in the input complex and the output complex.

1. In the Project Table, click the In column in row 1 to include only the input entry in the Workspace.
2. Click the Draw bonds in tube toolbar button



and click on an atom in the ligand molecule.

The ligand is displayed in the tube representation.

3. Choose Select from the Display only button menu.



The Atom Selection dialog box is displayed.

4. In the Molecule folder, choose Molecule List.
5. Select the ligand molecule, PGH, from the list.
6. Click Add.

7. Click the Proximity button.

The Proximity dialog box opens.

8. Under Proximity select Within and Angstroms, and enter 3 in the text box.
9. Under Fill select Residues.
10. Ensure that Exclude source is not selected.

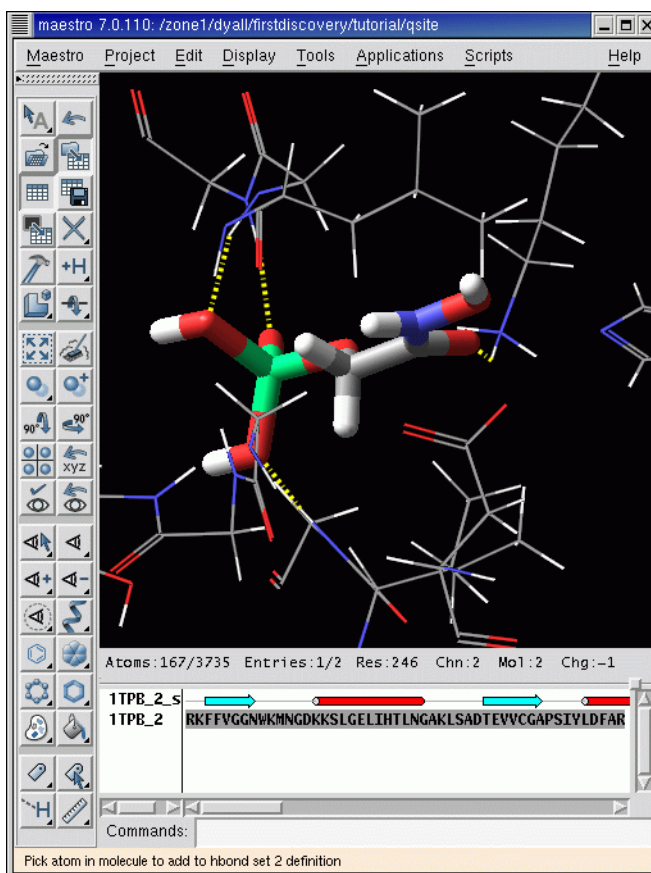


Figure 3.6. Input structure: ligand-receptor hydrogen bonds.

11. Click OK.

Residues with any atoms within 3 Å of the ligand are added to the display in wire frame.

12. Include only the output (second) entry in the Workspace, then repeat [Step 2](#) – [Step 11](#).

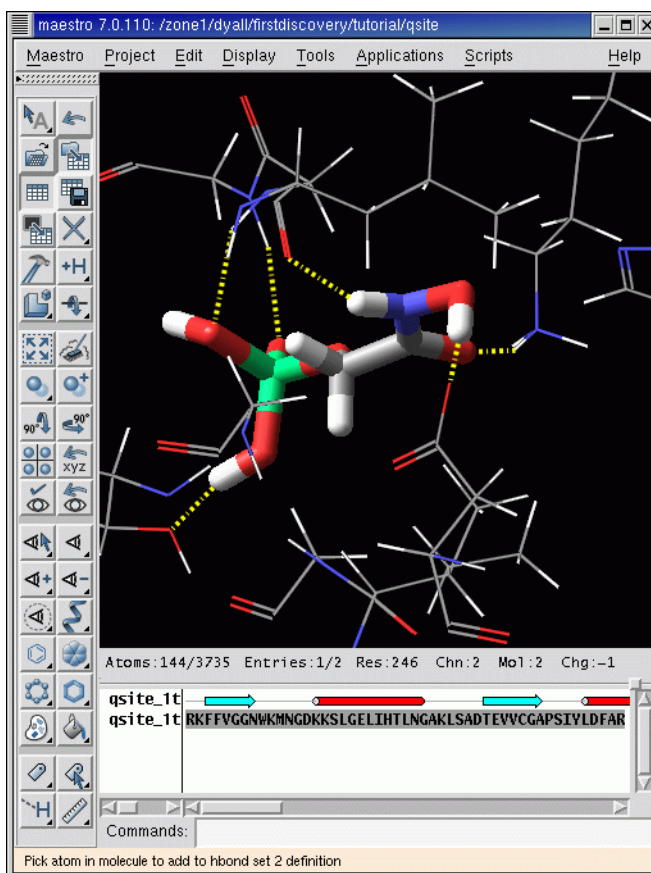
13. Include both entries in the Workspace.

14. Choose Atom Type (MacroModel) from the Color all atoms by scheme button menu.



Corresponding atoms in the two entries are now colored identically.

15. Include only the input entry.



**Figure 3.7. Output structure: ligand-receptor hydrogen bonds.**

16. Choose Measurements from the Tools menu.

The Measurements panel opens.

17. In the H-Bonds folder, under Atom set 1, use the picking controls to select the ligand molecule.
18. Under Atom set 2, select Pick, then use the picking controls to select the receptor molecule.

Hydrogen bonds between the ligand and nearby receptor residues are shown as dashed yellow lines, as shown in [Figure 3.6](#).

19. Control-click the In check box for the output entry. Both entries are now included.
20. Control-click the In check box for the input entry to exclude it.

Using control-click to include and exclude entries allows the Atom set 1 and Atom set 2 definitions in the H-Bonds folder of the Measurements panel to stay the same.

As you include the output structure and exclude the input structure, notice how the hydroxamate group rotates slightly to form hydrogen bonds with leucine (LEU 230) and aspartic acid (ASP 165) in the output structure ([Figure 3.7](#)).

# Protein Preparation

The quality of QSite results depends on reasonable starting structures. Schrödinger offers a comprehensive protein preparation facility designed to ensure chemical correctness and to optimize protein and protein-ligand complex structures for use as input. For QSite, the entire system (both MM and QM regions) must be prepared so as to satisfy the requirements of Impact/OPLS2001, while the QM region must satisfy the requirements of Jaguar. It is strongly recommended that protein structures imported from non-Maestro sources, such as PDB structures, be treated with the protein preparation facility in order to achieve best results.

For the molecular mechanics simulations using Impact with OPLS force fields, structures must be all-atom (explicit hydrogens) and there must be no covalent bonds between ligand atoms and protein atoms, including protein metal atoms. Bond orders and formal charges must be correct.

For Jaguar calculations on the QM region, structures must be three-dimensional and correct (have reasonable starting bond orders and formal charges).

This chapter describes the preparation of protein-ligand complexes. Several preparation tasks must be performed manually before using the protein preparation facility for the automated tasks. Most of the features of the facility are available from the Glide Protein Preparation panel. Additional features are available in the command-line application `protprep`. The utilities `pprep` and `impref` are also available. Use of the command-line application and utilities is summarized in [Section 4.10 on page 53](#).

## 4.1 The Protein Preparation Facility

The protein preparation facility performs the final stages of the preparation of proteins for use in QSite, Glide, CombiGlide, and Liaison. A typical PDB structure file consists only of heavy atoms; therefore, hydrogens have to be added prior to use in QSite calculations, which use an all-atom force field. The charge state of protein residues is also important to the results generated by QSite. Before you can run a protein preparation job, you must perform some preliminary preparation tasks that are not automated.

The protein preparation facility consists of two components, *preparation* and *refinement*. After ensuring chemical correctness, the *preparation* component adds hydrogens and neutralizes side chains that are not close to the binding cavity and do not participate in salt bridges. The *refinement* component performs a restrained Impact minimization of the cocrystallized

complex, which reorients side-chain hydroxyl groups and alleviates potential steric clashes. The Protein Preparation panel is used to set up jobs that perform these tasks.

If you are familiar with Maestro, or if you have followed Schrödinger's protein preparation procedure before, you might need to follow only the overview provided in [Section 4.2](#). The steps are described in detail in later sections of this chapter. For a tutorial on protein preparation, see [Chapter 3](#) of the *Glide Quick Start Guide*.

## 4.2 Step-by-Step Overview

This section provides an overview of the protein preparation process. The procedure described assumes that the initial protein structure is in a PDB-format file, includes a cocrystallized ligand, and does not include explicit hydrogens. For best results, structures with missing residues near the active site should be repaired before protein preparation. After processing with Schrödinger's protein preparation facility, you will have files containing refined, hydrogenated structures of the ligand and the ligand-receptor complex. The prepared structures are suitable for use with QSite. In most cases, not all of the steps outlined need to be performed. See the descriptions of each step to determine whether it is required.

1. Import a ligand/protein cocrystallized structure, typically from PDB, into Maestro. The preparation component of the protein preparation facility requires an identified ligand.
2. Locate any waters you want to keep, then delete all others.

These waters are identified by the oxygen atom, and usually do not have hydrogens attached. Generally, all waters (except those coordinated to metals) are deleted, but waters that bridge between the ligand and the protein are sometimes retained. If any waters are kept, hydrogens will be added to them by the preparation component of the protein preparation job. Afterwards, it is useful to check that these water molecules are correctly oriented.

3. Simplify multimeric complexes.
  - Determine whether the protein-ligand complex is a dimer or other multimer containing duplicate binding sites and duplicate chains that are redundant.
  - If the structure is a multimer with duplicate binding sites, remove redundant binding sites and the associated duplicate chains by picking and deleting molecules or chains in Maestro.
4. Adjust the protein, metal ions, and cofactors.
  - Fix any serious errors in the protein. Incomplete residues are the most common errors, but are relatively harmless if they are distant from the active site. Structures that are missing residues near the active site should be repaired.

- Check the protein structure for metal ions and cofactors.
- If there are bonds to metal ions, delete the bonds, then adjust the formal charges of the atoms that were attached to the metal as well as the metal itself.
- Set charges and correct atom types for any metal atoms, as needed.
- Set bond orders and formal charges for any cofactors, as needed.

5. Adjust the ligand bond orders and formal charges.

If you are working with a dimeric or large protein and two ligands exist in two active sites, the bond orders have to be corrected in both ligand structures.

6. Run protein preparation.

- Open the Protein Preparation panel, define the ligand by picking, select the desired Procedure, set other options as necessary, and click Start.

The Protein Preparation Start dialog box opens.

- Type the name of your job in the Name text box, and click Start.

7. Review the prepared structures.

- If problems arise during the preparation or refinement stages, review the log file, correct the problems, and rerun.
- Examine the refined ligand/protein/water structure for correct formal charges and protonation states resulting from Step 6 and make final adjustments as needed.

## 4.3 Importing the Protein Complex Structure

This step begins the protein preparation procedure.

### To import a ligand-receptor protein complex structure into Maestro:

1. On the toolbar, click the Import structures button:



The Import panel is displayed.

2. Select PDB format.
3. Enter the name of the file, or select the file in the Files list.
4. Click Import.

5. To display the Project Table, click the Open/Close project table button on the toolbar:



The imported entry is highlighted in the Project Table and displayed in the Workspace.

## 4.4 Deleting Unwanted Waters

Water molecules in the crystallographic complex are generally not used unless they are judged critical to the functioning of the protein–ligand interaction. When waters are used, they are later included in the protein as “structural” waters.

If one or more water molecules is believed to be involved in bond breaking or bond formation, it should be retained, preferably as part of the QM region. Though a single water molecule does not greatly increase the size of the SCF calculation, it can have considerable freedom of movement, which could delay or even prevent convergence of a geometry optimization or transition-state search. Water molecules that are not directly involved in ligand binding or enzymatic reactions can be deleted or included in the MM region.

You can delete all waters by choosing Waters from the Delete button menu:



## 4.5 Simplifying the Protein Complex

### 4.5.1 Determining Whether the Complex Is a Multimer

To determine whether the ligand-receptor complex is a multimer, compare the chains that appear in the sequence viewer. If there are two or more chains with identical sequences, the complex may be a multimer. If this is the case, there may be duplicate copies of the binding site of interest, with duplicate chains forming the duplicate binding sites.

If the binding interaction of interest takes place within a single subunit, you should retain only one ligand-receptor subunit to prepare for QSite. However, if two identical chains are both required to form the active site, neither should be deleted. To see whether two duplicate chains are involved with the active site, undisplay the protein’s amino acid residues:

1. On the toolbar, choose Protein Backbone from the Undisplay button menu:



2. Repeat the process and choose Protein Side Chains.

Ligands, cofactors, metal ions, and water oxygens remain visible. If two or more identical ligands or ligand/cofactor groups are present, then the complex is most likely a multimer, and the redundant groups and the duplicate chains associated with them can be deleted.

### 4.5.2 Retaining Needed Subunits

If the protein complex structure is a multimer with duplicate binding sites, it can be truncated by deleting all but a single ligand binding site and the associated receptor subunits. If you choose not to truncate the structure, skip to [Section 4.6](#). The process below assumes that the protein is initially not displayed.

**To remove redundant subunits or receptor sites of a multimer:**

1. Delete all but one ligand or ligand/cofactor pairing:
  - a. Choose Molecules from the Delete button menu.
2. Display the ligand or ligand/cofactor pair in CPK:
  - a. Choose Molecules from the Draw atoms in CPK button menu.



- b. Click on any atom in a molecule to delete that molecule.



- b. Click on an atom in the ligand to display it in CPK.
  - c. If there is a cofactor, click on an atom in that molecule as well.
  - d. Click the toolbar button a second time to leave the Draw atoms in CPK pick state.
3. Redisplay the protein backbone by choosing Protein Backbone from the Also display button menu.



Making just the backbone visible will provide enough information without unduly cluttering the Workspace.

4. Assign coloring by Chain Name by choosing Chain Name from the Color all atoms by scheme button menu.



5. Delete duplicate protein chains:
  - a. Choose Chains from the Delete button menu.
  - b. Click on a backbone atom in each protein chain you want to delete.
6. Delete duplicate ligands and cofactors:
  - a. Choose Molecules from the Delete button menu.
  - b. Click on an atom in each ligand or cofactor to be deleted.
7. When finished, redisplay the rest of the protein:

Choose All from the Display only button menu.



8. Redisplay the entire structure in wire-frame representation by double-clicking the Draw bonds in wire button.



## **4.6 Adjusting the Protein, Metal Ions, and Cofactors**

Problems originating in the PDB protein structure may need to be repaired before it can be used. Incomplete residues are the most common errors, but are relatively harmless if they are distant from the active site. Structures that are missing residues near the active site should be repaired.

If the protein already includes hydrogen atoms, you will need to decide how to proceed. If all hydrogens are present, you could use the structure as is and omit running the protein preparation procedure. This approach is not recommended unless you are absolutely satisfied that the structure is properly prepared and contains no untenable steric clashes.

Typically, you will need to perform the tasks in this section to assure that the protein structure is ready to be prepared.

### 4.6.1 Checking the Protein Structure for Metal Ions and Cofactors

1. Ensure that the entire structure, including any metals and cofactors, is included in the Workspace.
2. Choose Element from the Color all atoms by scheme button menu:



All atoms in the Workspace are now colored by element.

3. Choose Protein Backbone from the Undisplay button menu.



4. Choose Protein Side Chains from the Undisplay button menu.
5. As needed, redisplay protein residues near a metal or cofactor, using the Also display button menu and the Atom Selection dialog box.



6. Examine the protein structure to determine how to continue.
  - If there are formal bonds from protein atoms to metal ions, see [Section 4.6.2](#). Then correct the metal ions as in [Section 4.6.3](#).
  - If the protein contains metal ions without bonds, treat them as described in [Section 4.6.3](#).
  - If the protein contains cofactors, treat them as described in [Section 4.6.4](#).
  - Bonds to the ligand will be discussed in [Section 4.7](#).

### 4.6.2 Deleting Protein-Metal Bonds

Impact does not permit covalent bonds between metals and protein atoms. For metalloenzyme studies, the metal ion and adjacent protein side chains are best treated quantum mechanically; however, the QM region in QSite must meet Impact's requirements as well. The preparation stage of protein preparation automatically deletes protein-metal bonds before hydrogens are added, but this can leave incorrect formal charges on the atoms that were bonded to metal, causing incorrect structures to be generated by protein preparation.

It is safest, therefore, to delete the bonds, then check and adjust element names and formal charges for both metal and non-metal atoms whose bond order you have changed. When you

have run protein preparation, it is highly recommended that you examine the prepared complex for correct protonation states and charges in the active site, then make manual corrections if needed. Before running a QSite job, you can run a score-in-place Glide calculation on the complex and check that the metal-ligation energy is reasonable. If it is highly positive, you may have to re-adjust the charge and protonation states in the active site manually.

### To manually delete bonds between metals and protein atoms:

1. Choose Bonds from the Delete button menu:



2. Click on the bonds to be deleted.

### 4.6.3 Adjusting Metal Ion Charges

The MacroModel atom types for metal ions are sometimes incorrectly translated into dummy atom types (Du, Z0, or 00) when metal-protein bonds are specified in the input structure. Furthermore, isolated metal ions may erroneously be assigned general atom types (GA, GB, GC, etc.). The protein preparation procedure cannot treat structure files containing these atom types; they should be corrected as described in this section.

### To display element labels and formal charges:

1. Open the Build panel by clicking the Open/Close build panel button:



2. In the Build panel toolbar, click the Label button:



All metal ions (and other heteroatoms) are labeled with their element symbol and formal charge.

3. Check any metal ions to make sure they are correct. If they are, the next step in the process is [Section 4.6.4](#). If not, you can correct them.

### To correct metal ion atom types:

1. In the Build panel, click the Atom Properties tab and select Atom Type (MacroModel) from the Property option menu.

2. Select the correct atom type for the metal ion from the list.

The atom type for metal ions includes both element name and formal charge. Atom type numbers are in parentheses.

3. Click on the metal ion to be changed.

#### 4.6.4 Displaying and Adjusting the Cofactor

Cofactors are included as part of the protein, but because they are not standard residues it is sometimes necessary to use Maestro's structure-editing capabilities to ensure that multiple bonds and formal charges are assigned correctly.

1. If the Build panel is not displayed, click Open/Close build panel on the main toolbar.



##### To display only the cofactor:

1. Choose Select from the Display only button menu.



The Atom Selection dialog box is displayed.

2. In the Residue folder, choose Residue Type.
3. Click the residue type of the cofactor, which will be near the end of the list.

The cofactor is highlighted.

4. Click Add, then click OK.

The cofactor is displayed. Because the cofactor was chosen by residue type and not molecule number, this method works even if the cofactor is covalently bonded to another residue.

##### To set or change cofactor bond orders:

1. On the Build panel toolbar, click the Decrement bond order or Increment bond order button, as appropriate.



2. Click on bonds as necessary to set the bond order.

**To set or correct the formal charge on any cofactor atoms:**

1. On the Build panel toolbar, click the Label button.



All metal ions (and other heteroatoms) are labeled with their element symbol and formal charge.

2. On the Build panel toolbar, click on the Increment formal charge or Decrement formal charge button, as appropriate.



3. Click on an atom whose formal charge must be increased or decreased. Repeat as necessary. The atom labels show the current formal charge.

**To correct the atom type of any mistyped atoms:**

1. In the Atom Properties folder of the Build panel, choose Atom Type (MacroModel) from the Property option menu.
2. Select the correct atom type for the mis-typed atom from the list.
3. Click on the atom to be changed.
4. If the cofactor contains any metal ions, bonds between the metal and cofactor can be removed as in [Section 4.6.2](#).

For more information about structure editing in Maestro, click Help or see [Chapter 4](#) of the *Maestro User Manual*.

## 4.7 Adjusting the Ligand

### 4.7.1 Manually Deleting Protein-Ligand Bonds

The use of Impact in protein preparation (restrained minimization during refinement) prohibits the existence of formal protein-ligand bonds. This category includes bonds between a protein metal atom and a ligand or cofactor. The OPLS-AA force field models such interactions as a van der Waals plus electrostatic interaction, and the presence of explicit covalent bonds to the ligand, as might be found in an acyl enzyme, interferes with this model. Therefore, Maestro deletes all bonds between the ligand and metal atoms before the protein preparation job begins.

However, setting up the protein preparation job requires the identification of the ligand molecule, and this identification is disrupted by the existence of covalent bonds between protein atoms (metals or otherwise) and ligand atoms. To avoid this problem, you should delete such bonds manually.

1. To check for protein-ligand bonds, including protein metal-ligand bonds, use Maestro's Display/Undisplay facility to undisplay the protein backbone and protein side chains. The ligand and any cofactors and metals are displayed.
2. If any metal-ligand or other protein-ligand bonds exist, delete them:
  - a. Choose Bonds from the Delete button menu on the toolbar.



- b. Click on the bonds to be deleted.
3. Redisplay the entire protein by choosing All from the Display Only toolbar button menu.



## 4.7.2 Adjusting Ligand Atom and Bond Properties

You will need to adjust the formal charges of any atoms, metal or otherwise, from which you have deleted a bond. Follow the procedures for adjusting metal and cofactor atom and bond properties, as described in [Section 4.6.3](#) and [Section 4.6.4](#).

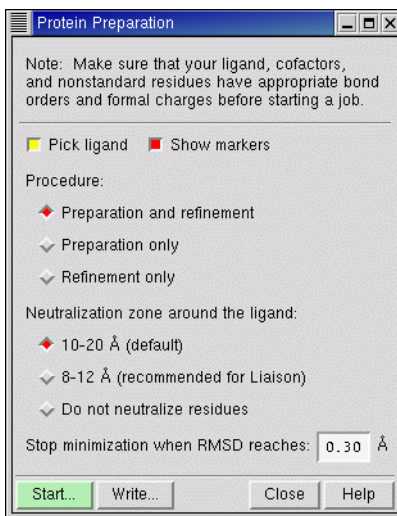
## 4.8 Running Protein Preparation on the Structures

From this point on, all structural manipulations are done by the Protein Preparation panel, shown in [Figure 4.1](#), and its related scripts. Before you open this panel, ensure that the protein and ligand are in the Workspace.

To open the Protein Preparation panel, select Protein Preparation from the Glide submenu of the Applications menu on the main menu bar.

### 4.8.1 Defining the Ligand

Before launching a protein preparation job, you must choose a molecule in the Workspace that will be treated as the ligand. In the Protein Preparation panel, select Pick ligand, and then select the ligand by clicking on it in the Workspace. When Show markers is selected, the ligand is highlighted with a blue-green marker. The rest of the Workspace is then treated as the protein.



**Figure 4.1. The Protein Preparation panel.**

### 4.8.2 Choosing a Procedure

The Protein Preparation panel facilitates three types of jobs: Preparation only, Refinement only, and Preparation and refinement.

The Preparation component neutralizes residues that are beyond a set distance from the ligand. The Preparation process also detects some conflicts in hydrogen bonding. It corrects them when possible, either by exchanging carbonyl and hydroxyl oxygens in a neutralized carboxylic acid group, or by creating the alternate (HIE) tautomer of a histidine side chain.

The Refinement component uses Impact to run a series of restrained, partial minimizations on the combined, hydrogenated structure. Minimizations continue until the average RMS deviation of the non-hydrogen atoms reaches the specified limit (0.3 Å by default).

The first step in the sequence of restrained minimizations reorients side-chain hydroxyls in serine, threonine, and tyrosine residues, and side-chain sulfhydryls of cysteine residues. This is accomplished by tightly tethering non-hydrogen atoms with a force constant of 10 kcal/mol·Å<sup>2</sup> and by minimizing the hydrogens with torsion interactions turned off.

Each restrained minimization employs a limited number of minimization steps and is not intended to minimize the system completely. Subsequent steps restore the torsion potential and use progressively weaker restraints on the non-hydrogen atoms (hydrogen atoms are always free). The force constants employed are 3, 1, 0.3, and 0.1 kcal/mol·Å<sup>2</sup>.

Preparation and refinement, the default, runs both components. This is the recommended mode if you have not yet run any preparation jobs on the protein. Separate Preparation only and Refinement only jobs can be run if you encountered a problem in the combined Preparation and refinement job. Subsequent Refinement only jobs can be performed after a Preparation and refinement job if water molecules need to be reoriented or if other structural adjustments need to be made.

### 4.8.3 Other Options

1. Neutralization zone around the ligand

The default is 10-20 Å. You can choose a shorter distance (8-12 Å) or choose Do not neutralize residues.

2. Stop minimization when RMSD reaches: 0.30 Å

This is the default value. It allows the refinement portion of the job to halt when the average RMSD of the heavy atoms reaches 0.30 Å.

### 4.8.4 Starting the Protein Preparation Job

To start the protein preparation job, click Start. The Start dialog box is displayed, in which you can select a job name and a host, and enter your user name on the host if it is different from the host on which Maestro is running.

If you decide to run preparation and refinement separately, you will need to run a Preparation only job and refine the results with a subsequent Refinement only job.

If you want to run the job from the command line, click Write to write the input files you will need without starting the job. See [Section 4.10 on page 53](#) for information on command-line options.

Click Start in the Start dialog box when you are ready to launch the job. When the job starts, the monitor panel is displayed, including the running log of the job. When the job finishes, the results are appended to the Project Table. The protein and the ligand are incorporated as separate entries.

### 4.8.5 Output Job Files

Running Preparation and refinement produces the same files as running Preparation only followed by Refinement only. The following structure files are produced, where *struct* is the name of the complex:

<code>struct_lig.mae</code>	The input ligand structure file
<code>struct_lig_prep.mae</code>	The post-preparation ligand structure file
<code>struct_lig_ref.mae</code>	The post-refinement ligand structure file
<code>struct_prot.mae</code>	The input receptor structure file
<code>struct_prot_prep.mae</code>	The post-preparation receptor structure file
<code>struct_prot_ref.mae</code>	The post-refinement receptor and ligand structures
<code>struct.log</code>	The log file for the complete preparation and refinement job

## 4.9 Checking the Output Structures

Finally, after both the preparation and refinement components have successfully run, you should check the completed ligand and protein structures.

### 4.9.1 Checking the Orientation of Water Molecules

Perform this step only if you identified and kept some structural waters in [Section 4.4](#). Reorienting the hydrogens is not strictly necessary, as their orientation should have been changed during refinement in [Section 4.8](#), but it is useful to check that the orientation is correct.

If the orientation is incorrect, reorient the molecules by using the following procedure:

1. Choose Global/Local from the Local transformation button menu.



The Advanced Transformations panel is displayed.

2. Under Atoms For Transformation, use the picking controls to select the entire water molecule you want to reorient.
3. Under A Center For Transformation, use the picking controls to select the oxygen atom of the water molecule.
4. Under Rotation/Translation Scope, select Local.
5. Use the middle mouse button to change the orientation of the water hydrogens.
6. Close the Advanced Transformations panel. Transformations should now be global again.

When you have corrected the orientation of the retained water molecules, run a Refinement only job on the adjusted protein-ligand complex as described in [Section 4.8](#).

### 4.9.2 Resolving H-Bonding Conflicts

One or more residues may need to be modified to resolve an acceptor-acceptor or donor-donor clash. If residues need to be modified, follow these steps:

1. Place the refined protein-ligand complex in the Workspace.
2. Examine the interaction between the ligand and the protein (and/or the cofactor).
3. Use your judgment and chemical intuition to determine which protonation state and tautomeric form the residues in question should have.
4. Use the structure-editing capabilities in Maestro to resolve the conflict.

Some of these clashes are recognized by the preparation process but cannot be resolved by it. The preparation process may have no control over other clashes. An example of the latter typically occurs in an aspartyl protease such as HIV, where both active-site aspartates are close to one or more atoms of a properly docked ligand. Because these contact distances fall within any reasonable cavity radius, the carboxylates are not subject to being neutralized and will both be represented as negatively charged by the preparation process. However, when the ligand interacts with the aspartates via a hydroxyl group or similar neutral functionality, one of the aspartates is typically modeled as neutral.

## 4.10 Command-Line Protein Preparation

**To run protein preparation from the command line:**

1. If you do not yet have receptor and ligand structure files for the structures in the Workspace, use the Write button in the Protein Preparation panel to write the structure files.
2. Use the `protprep`, `pprep`, or `impref` command-line utilities to run specific procedures. These commands and their options are summarized below.

### 4.10.1 Usage Summary for `protprep`

The `$$SCHRODINGER/protprep` application has command-line options corresponding to features of the Maestro Protein Preparation panel. The command `protprep -h` displays the usage summary that appears in this section.

**Syntax:**

```
$$SCHRODINGER/protprep [options] input-file
```

*input-file* is the file containing the protein to be prepared or refined. This file must be in Maestro format. When doing a refinement only job (`-mode refine`) this file can contain a protein-ligand complex.

### Options:

#### General Options:

-j <i>jobname</i>	Override the default job name derived from <i>input-file</i> . This allows you to choose an output job name that is different from the <i>input-file</i> name.
-l <i>ligand-file</i>	Specify a file containing a ligand in the protein's active site. This file must be in Maestro format.
-m <i>mode</i>	Select mode, where <i>mode</i> is one of the following:
-mode <i>mode</i>	prep      Preparation only. refine    Refinement only. both      Preparation and refinement (default).
-debug	Print verbose (debugging) output.
-HOST <i>host</i>	Run the job on a remote host.
-LOCAL	Run the job in the current directory, rather than in a temporary scratch directory.
-WAIT	Wait until the job finishes before returning the command prompt.
-NICE	Run the job at reduced priority.
-HELP   -h	Print usage message and exit.

#### Preparation Stage Options:

-min-recep-only	Minimize total charge of receptor only.
-skip-sidechain-corr	Skip correction of conflicting side-chain forms.
-cavity-8-12	Set cavity distance range to 8-12 Å. Suitable for Liaison jobs.
-salt-bridge-dist	Leave residue pairs forming salt bridges within <i>dist</i> Å ionized; default is 3.5 Å.
-ionization-range	Leave residues within <i>dist</i> Å of ligand ionized.
-hbond-dist	Set H-bonding distance; default is 3.45 Å.

#### Refinement Stage Options:

-r <i>rmsd</i>	Maximum RMSD allowed for refinement; default 0.3.
-keep	Keep intermediate structure files.
-separate	Write out refined protein and ligand structures separately, rather than in one combined structure.

### 4.10.2 Usage Summary for **pprep**

The purpose of **pprep** is to adjust protonation states of a receptor in a Maestro format file. **pprep** is the driver for the preparation stage, and is called by **protprep**. There should be little need to run **pprep** directly.

#### Syntax:

```
$SCHRODINGER/utilities/pprep [options] proteinfile.mae
```

#### Options:

- i *idis*            Leave residue pairs forming salt bridges within *idis* ionized; default is 3.5 Å.
- l *ligfile*        Read ligand mae file *ligfile*.
- n *outfile*        Specify non-default (*ligR.mae*) name for output file with neutralized residues.
- p                Print verbose output.
- r                Minimize total charge of the receptor only.
- t                Skip correction of conflicting side-chain forms.
- w *wdis*           Leave residues within *wdis* of ligand ionized.
- H *hbonddist*    Set H-bonding distance; default 3.45 Å.
- L                Set cavity distance range to 8-12 Å.
- v                Print version number and exit.
- h                Print usage message and exit.

### 4.10.3 Usage Summary for **impref**

The purpose of **impref** is to use Impact for restrained optimizations of a ligand-receptor complex. **impref** is the driver for the refinement stage, and is invoked by **protprep**. There is little need to run **impref** directly.

#### Syntax:

```
$SCHRODINGER/utilities/impref [options] input.mae
```

**Options:**

- k Keep Impact minimization \*.inp, \*.log, and \*.mae files.
- l *ligfile* Read ligand from file *ligfile*, instead of *input.mae*.  
If this option is used, *input.mae* must be the protein structure alone.  
If this option is not used, *input.mae* must be the protein/ligand complex.
- r *rmsd* Specify maximum RMSD allowed; default is 0.3.
- s Write out protein and ligand separately. Requires -l *ligfile*.
- op *file* Output protein or complex file. Default is *input\_ref.mae*.
- ol *file* Output ligand file (when -s and -l used.) Default is *ligfile\_ref.mae*.
- v Print version number and exit.
- h Print usage message and exit.

# Running QSite From Maestro

QSite performs mixed quantum mechanical/molecular mechanical (QM/MM) calculations, using Jaguar for the QM calculations and Impact for the MM calculations. Ligands and other specified regions of a protein complex can be studied using QM, while MM is used for the rest of the molecule.

At each step of a QM geometry optimization, Impact calculates energy terms for MM-QM region interactions; if MM minimization was also specified, it is also performed at each QM step. The next QM step takes into account the new MM atom distribution and energy terms. If a single-point QM calculation is selected, the current QM/MM energy is calculated without MM minimization.

The speed of QSite is largely determined by the size of the QM region. Therefore there is no advantage to making a smaller model protein. You can run calculations on systems with up to than 8000 atoms or 8000 bonds from the QSite panel, but larger systems must be run them from the command line. See [Section 5.7 on page 72](#) for more information.

Cartesian constraints may be placed on atoms in both the QM and the MM regions. See [Section 5.3 on page 62](#) for a description of the two types of constraints. Frozen-atom constraints can be applied to atoms in both regions. Constrained atoms can be specified for MM-region atoms, but are ignored if applied to QM-region atoms.

In general, a QSite calculation can only be performed using a single entry. If you want to run a QSite job using the Workspace structure as input, and that structure includes multiple entries, combine them into a single entry using the Merge option from the Entry menu in the Project Table panel. The merged entry should be the only entry included in the Workspace when you start the job. One exception to this is when setting up a transition-state search. In this case you may select up to three Project Table entries, depending upon the algorithm that is selected for performing the search. See [Section 5.5](#) for more information about transition-state searching.

## 5.1 The QSite Panel

QSite calculations can be set up and run using the QSite panel. To open the QSite panel, choose QSite from the Applications menu.

At the top of the panel is an option menu for selecting the source of structure input, which is described in the next subsection. The center section of the panel has five tabbed folders:

- Potential—choose settings for the MM potential energy function.
- Constraints—set up atom constraints for atoms in the MM and QM regions.
- Minimization—set up energy minimization of the MM region.
- Optimization—select and set up the job task.
- QM Settings—specify the QM region and other QM options.

These folders are described in later sections of this chapter. Below the folders are several action buttons, which are described in a subsection below.

### **5.1.1 Source of Structure Input**

Input for QSite single point jobs, minimization jobs, and standard transition state searches must be a single Project Table entry. To use a system consisting of two or more entries as the input, choose Merge from the Entry menu in the Project Table panel, create a combined entry, and run the simulation on that entry. LST and QST transition state searches are a special case: here, the input is specified in the Optimization folder.

You can select the source of input from the Use structures from option menu. The options are described below.

#### **Workspace (included entry)**

This is the default option. The structure that is currently included in the Workspace is used as input to the job. This includes whatever atoms, molecules, or entries are part of the structure, even atoms that have been undisplayed. You must choose this option if you want to use constraints (see the Constraints folder, [Section 5.3 on page 62](#)), because the constraints are applied to the Workspace structure.

#### **Project Table (selected entry)**

Select this option to use the entry that is currently selected in the Project Table. This may be different from the structure in the Workspace. Because atom constraints are applied to the Workspace structure, they are ignored if this option is chosen.

### **5.1.2 Action Buttons**

The lower part of the QSite panels contains a row of action buttons. Apart from the standard Close and Help buttons, for closing the panel and opening the help viewer, there are two buttons for job-related actions, described below.

#### **Start**

Click the Start button to open the Start dialog box. [Section 5.7](#) describes how to start jobs using the Start dialog box.

## Write

The Write button writes out all the files required for the job but does not run the job. Once the files required for the job are written by Maestro, the job can be run from the command line in a terminal window using the syntax:

```
$SCHRODINGER/qsite job-options jobname
```

where *jobname.in* is the input file for the job in question, and *job-options* is a list of options for the job. The log output is written to *jobname.log*.

Type `$SCHRODINGER/qsite` for a usage summary of the `qsite` command, or see [Chapter 6](#) for information on running QSite from the command line.

## 5.2 The Potential Folder

The Potential folder provides options for the definition of the potential energy functions used in the molecular mechanics part of the calculation. One of these (continuum solvation) affects the QM potential energy as well.

### Force field

The default OPLS-AA force field used in QSite is OPLS\_2001, although the OPLS\_1999 force field is also available. However, QSite has been re-parametrized, and new backbone and side-chain cuts have been added, for OPLS\_2001 only; therefore this force field is recommended.

The parameter file `paramstd.dat` is provided in the QSite distribution and should not be changed.

### Electrostatic treatment

This option menu offers two methods for calculating the electrostatic component of the molecular mechanics energy:

- Constant dielectric

This option calculates the electrostatic interaction between atoms *i* and *j* as:

$$E_{\text{ele}} = 332.063762 \, q_i q_j / (\epsilon \, r_{ij})$$

A constant dielectric is appropriate for a vacuum (gas-phase) calculation or when an explicit or implicit solvent model is used.

- Distance dependent dielectric

This option calculates the electrostatic interaction between atoms  $i$  and  $j$  as:

$$E_{\text{ele}} = 332.063762 q_i q_j \epsilon r_{ij}^{-2}$$

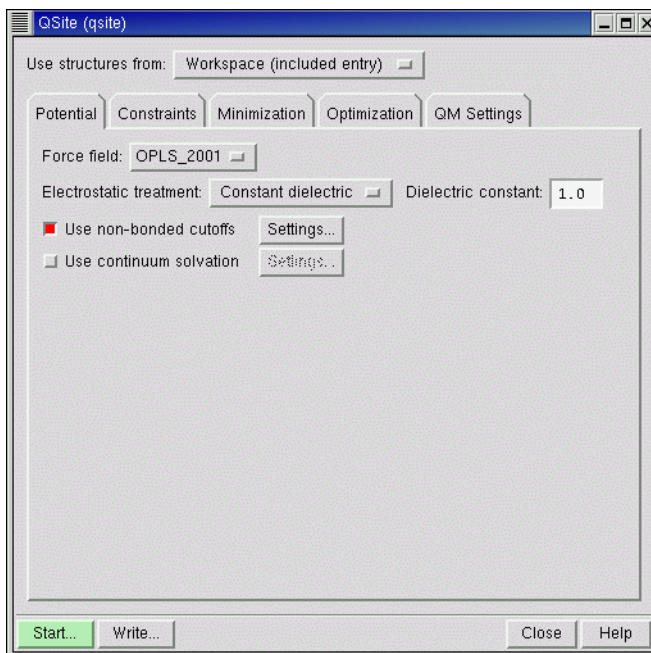
A distance-dependent dielectric is sometimes used as a primitive model for the effect of solvent. In this model, the electrostatic interaction between a pair of atoms falls off rapidly as the distance between the atoms increases. However, continuum and explicit solvent models are much better at accounting for solvent effects than a distance-dependent dielectric.

The variables in the above formulae are defined as follows:

- $E_{\text{ele}}$  is the electrostatic interaction in kcal/mol
- $q_i$  and  $q_j$  are the partial atomic charges on atom  $i$  and  $j$
- $r_{ij}$  is the distance in Å between atoms  $i$  and  $j$
- $\epsilon$  is the Dielectric constant (see below)

#### Dielectric constant

This text box specifies the value of the dielectric constant  $\epsilon$  used in the electrostatic calculations.



**Figure 5.1. The Potential folder of the QSite panel.**

### Use non-bonded cutoffs

In molecular-mechanics calculations it is often impractical to include the nonbonded (electrostatic and van der Waals) interactions between every pair of atoms. For large systems, many such pairs are separated by a great distance and contribute little to the interaction energy. Judicious truncation of the non-bonded interactions between widely separated pairs of atoms is an important strategy for reducing the resources needed for calculations on large systems.

At present only residue-based cutoffs are supported for calculations set up in Maestro. This means that all atoms within complete residues that have any pair of atoms within the cutoff distance will be included in the non-bonded interaction list. The list is updated periodically as the geometry changes, because residues may move inside or beyond the cutoff radius.

Select this option to truncate nonbonded interactions. Click **Settings** to open the Truncation panel. There are two settings that can be changed:

- **Update neighbor-list frequency (# steps):** The number of steps after which the neighbor list will be updated. The default is 10 steps (in the MM part of the calculation). Larger values will speed up the calculation, at the possible expense of accuracy.
- **Residue-based cutoff distance:** All atoms in complete residues that have any pair of atoms within this distance are included in the nonbonded interaction list. The default is 12 Å. This value should be increased to avoid convergence problems, to a value like 30 Å. Smaller values will speed up the calculations, but could miss important long-range electrostatic interactions between formally charged atoms.

### Use continuum solvation

This option affects both MM and QM calculations. Do not select Use continuum solvation if you intend to run the QM (Jaguar) calculation using multiple processors (parallel processing); when QSite jobs with solvation are run in parallel, erroneous energies result.

If you want to use explicit water solvent rather than continuum solvation, you can add the water molecules using the Soak application of Basic Impact, followed by an equilibration, also using Basic Impact. See [Chapter 6](#) of the *Impact User Manual* for more information on Soak.

Select this option and click **Settings** to open the Continuum Solvation dialog box. The only available solvation method in QSite, for both the MM and the QM solvation functions, is the Poisson–Boltzmann Solver (PBF); selecting Use continuum solvation automatically sets both solvation functions to PBF (in the Jaguar input file, `isolv=2`). The Surface Generalized Born (SGB) and Analytic Generalized Born (AGB) methods are not available for QSite calculations. The Continuum Solvation panel options for PBF are as follows:

- PBF resolution

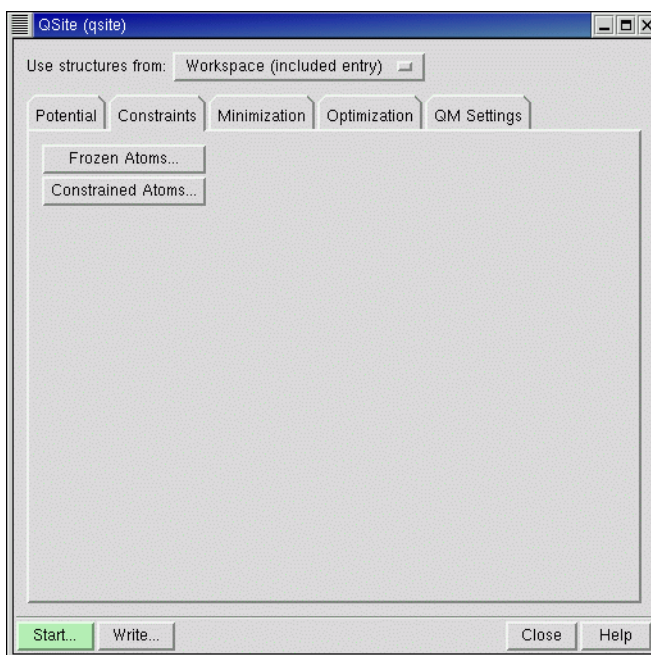
The Poisson–Boltzmann solver involves a finite-element calculation on a grid. The grid spacing controls the accuracy of the PBF calculation and the time required. The default, Low resolution, suffices for most protein work. If needed, greater accuracy can be achieved by choosing Medium or High resolution.

- PBF displacement threshold

This text box specifies how far (in Å) any atom may move from the coordinates used in the previous PBF calculation before a new PBF calculation must be performed. If no atom has moved this distance, the previously calculated PBF energy and forces are used.

## 5.3 The Constraints Folder

The Constraints folder is used to apply constraints to the Cartesian coordinates of selected atoms in the MM and the QM region. Specified atoms can be frozen at their input coordinates (frozen-atom constraints), or they can be constrained to remain near their initial coordinates by applying a harmonic force. In QSite, frozen-atom constraints can be applied to atoms in both the QM and the MM regions. However, only atoms in the MM region can be constrained; if QM-region atoms are selected to be constrained, the constraints are ignored.



**Figure 5.2.** The Constraints *folder of the* QSite *panel.*

Atom constraints in QSite, for atoms in the QM as well as the MM region, must be set using the Constraints folder. They *cannot* be set using a **coord** section in the *jobname.in* file.

The Constraints folder contains two buttons:

- Frozen Atoms
- Constrained Atoms

These buttons open the Frozen Atoms and Constrained Atoms panels. These panels have a similar structure. The panel features are described below.

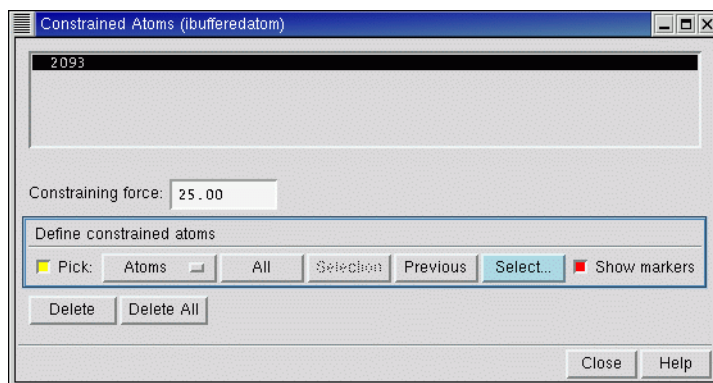
### Atoms list

The upper portion of each panel is a text area that lists the atom number of each atom that has been selected to be frozen or constrained. The currently selected atom is highlighted.

### Picking tools

This section, labeled Define Frozen Atoms or Define Constrained Atoms, contains standard picking controls: a Pick check box and menu, which is set to Atoms by default; an All button; a Select button, which opens the Atom Selection dialog box. If you deselect the Pick check box, you can use the Workspace selection tool to select atoms, then click the Selection button to add those atoms to the list, or use the Previous button to return to the most recent selection.

The picking tools section also contains a Show markers option, which is selected by default. Atoms to be frozen are marked with a red cross and a “padlock” icon in the Workspace. Atoms to be constrained are marked with a brown cross and a “spring” icon in the Workspace display. To distinguish the currently selected frozen or constrained atom, Maestro colors the marker turquoise.



**Figure 5.3.** The Constrained Atoms panel.

### Constraining force

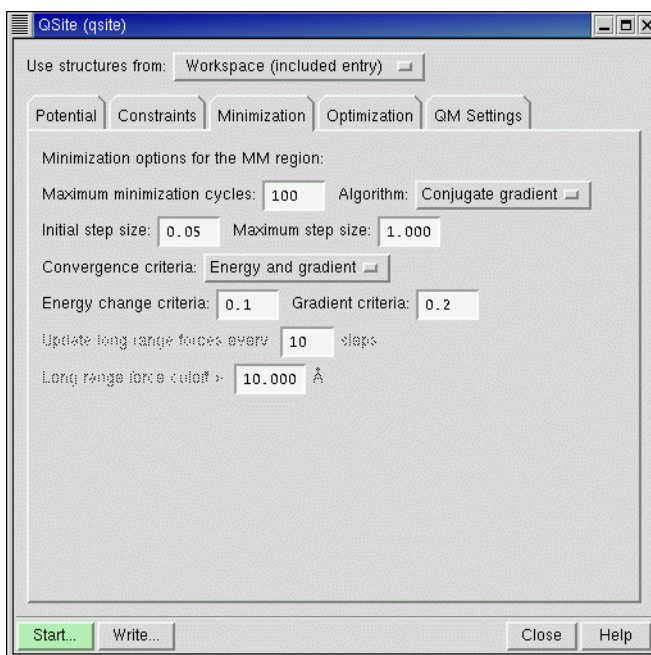
The Constraining force text box sets the value of the harmonic force constant applied to the selected constrained atoms. The same force constant is used for all atoms. The default is 25.00 kcal/(mol Å<sup>2</sup>).

### Deletion buttons

The constraint on the currently selected atom can be removed by clicking Delete. The atom is then removed from the atoms list. To remove all constraints of this type, click Delete All. The atoms list is then cleared.

## 5.4 The Minimization Folder

The Minimization folder specifies settings for Impact energy minimization of the MM region of the molecule. If the QM method chosen in the Optimization folder is Single point, these settings are not used, and no MM minimization is performed. The options available in this folder are described below.



**Figure 5.4.** The Minimization *folder of the* QSite *panel.*

**Maximum minimization cycles**

This text box sets the maximum number of cycles for the minimization calculation. The minimization terminates if it has not converged by this point. The default value of this setting is 100 iterations, but you can specify any value greater than or equal to zero. “Zero cycles” is a special case; it instructs Impact just to evaluate the energy for the current coordinates.

**Algorithm**

This option menu selects the minimization algorithm. The choices are:

- **Truncated Newton (TN).** This is a very efficient method for producing optimized structures. A short conjugate gradient pre-minimization stage is performed first to help improve the convergence of the Truncated Newton algorithm.
- **Conjugate gradient.** This is a good general optimization method and is the default method.
- **Steepest descent.** This can be a good method for initiating a minimization on a starting geometry that contains large steric clashes. Convergence is very poor towards the end of minimization, where the conjugate gradient algorithm should be used.

**Initial step size**

This text box specifies the initial step size of the minimization cycle for conjugate gradient and steepest descent minimizations. The default value is 0.05 Å, but any positive value is allowed.

**Maximum step size**

This text box specifies the maximum step size of the minimization cycle for conjugate gradient and steepest descent minimizations. If the step size exceeds this value, the minimization will halt. The default value is 1.00 Å, but any positive value is allowed. The maximum step size is the maximum displacement allowed for an atom in any step of a minimization calculation.

**Convergence criteria**

This option menu sets the convergence criteria for the minimization. Either or both of two criteria—energy change and gradient—can be specified. Thus, the options are:

- **Energy and Gradient.** Choosing this option allows access to both the Energy change criteria and Gradient criteria text boxes.
- **Energy change criteria.** Use this text box to specify the value of the energy change criterion. The default value is  $10^{-7}$  kcal/mol, but any positive value is allowed. The criterion is satisfied if two successive energies differ by less than the specified value.

- **Gradient criteria.** Use this text box to specify the value of the gradient criterion. The default value is 0.01 kcal/(mol Å), but any positive value is allowed. The criterion is satisfied if the norms of two successive gradients differ by less than the specified value.

**Long Range Forces Options (for Truncated Newton minimizations):**

- **Update long range forces every  $n$  steps.** Use this text box, when Truncated Newton minimization is selected, to specify the frequency with which long range forces are updated. Between these intervals, estimates of these forces are used. Every 10 steps is the default; smaller numbers (more frequent updates) can be used to improve convergence, but will make the optimization slower. Larger numbers for  $n$  may speed the calculation, but the maximum recommended value is 20.
- **Long range force cutoff  $> d$  Angstroms.** Use this text box, when Truncated Newton minimization is selected, to specify the distance beyond which forces are considered long range and are therefore updated every  $n$  steps. The default value is 10.000 Å.

## 5.5 The Optimization Folder

The Optimization folder specifies the type of QM (Jaguar) calculation to be performed and provides information needed to set up the calculation. For transition state optimizations additional structures (reactant, product, and transition state guess structures) can be given to guide the search. QSite geometry optimizations use internal coordinates by default; to force Cartesian coordinate optimization, run QSite from the command line, using the option `intopt=0` in the **gen** section of the Jaguar input file. The options available in this folder are described in the following sections.

### 5.5.1 Calculation Type

The Method menu controls the QM calculation type. The options are:

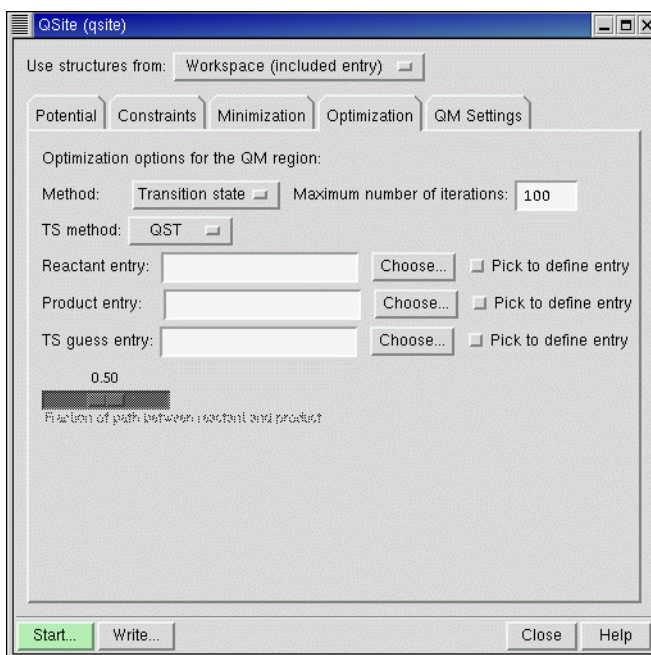
- **Single point**—Calculate the QM energy for the structure as it stands. No QM geometry optimization or MM minimization is performed. When Single point is selected, other options in this folder are unavailable. Settings in the Minimization and Constraints folders are ignored. This is the default QM method.
- **Minimization**—Locate a minimum-energy structure by geometry optimization. If you want to optimize only the QM region, simply set the number of minimization steps to 0 in the Minimization folder. There is no need to explicitly freeze all of the MM atoms.
- **Transition state**—Locate a transition state structure by geometry optimization. The remaining controls in the folder define the initial guess for the transition state geometry, and are described in the next subsection.

For minimization and transition-state calculations, you can specify the number of optimization iterations in the Maximum number of iterations text box. The default is 100 iterations.

### 5.5.2 Transition State Searches

If you select Transition state from the Method menu, the default option for TS method is Standard. The following three methods for transition-state optimization are supported in QSite, corresponding to well-known ab initio techniques. See [Section 5.3](#) of the *Jaguar User Manual* for detailed information about these methods:

- **Standard:** The standard transition-state optimization method is useful if you have only a single initial guess structure (the structure in the Workspace) for the transition-state. It attempts to find the saddle point closest to the starting structure by maximizing the energy along the lowest-frequency mode of the Hessian and minimizing the energy along all other modes.
- **LST:** *Linear Synchronous Transit* is useful if you have initial guess structures for the reactant and the product and want QSite to look for a transition-state structure by interpolating between them. LST uses a quasi-Newton method to search for the optimum transition-state geometry, choosing a transition-state guess structure based on the interpo-



**Figure 5.5.** The Optimization *folder* with TS method: QST *selected*

lation value you set using the Fraction of path between reactant and product slider. By default it is set at 0.50, directing QSite to choose an interpolated transition-state guess structure midway between the reactant and the product. If you want to pick a guess structure closer to the reactant, move this slider toward 0.00. For a guess structure closer to the product, move the slider toward 1.00

- **QST:** *Quadratic Synchronous Transit* is useful if you have initial guess structures for the reactant, the product, and the transition state. QST uses a quasi-Newton method to optimize the transition-state geometry.

The structures that define the transition state initial guess can be specified in the Reactant entry, Product entry, and TS guess entry sections. You can select the entry by typing in the entry name from the Project Table, by clicking Choose and selecting the entry from a list, or, if the structure is in the Workspace, by selecting Pick to define entry and clicking on any atom in the structure.

## 5.6 The QM Settings Folder

The QM Settings folder is used to enter information for the QM job and to define the QM region.

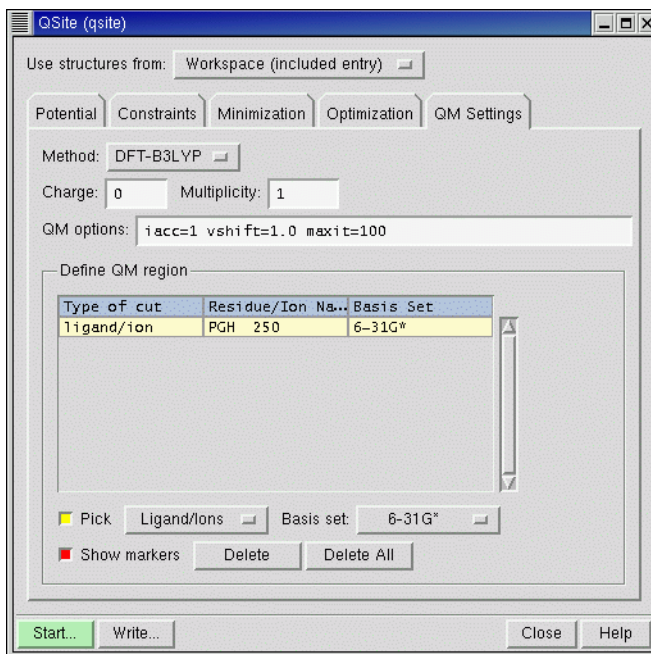
QM job information includes the quantum-mechanical method to be used, the charge and spin multiplicity of the QM system, the number of processors (if Jaguar parallel processing is available), and other keywords and options that may be required by Jaguar.

The QM region can be defined by either:

- Selecting the ligand, metal ions, or other disconnected species (not covalently bonded to the protein).
- Specifying *cuts* between certain covalently-bonded atoms in connected peptide residues. QSite cuts are specially parameterized frozen-orbital boundaries between the QM and MM regions. They can be placed between an alpha carbon and a side chain (side-chain cuts) or between an alpha carbon and the backbone to one side (backbone cuts, which must be made in pairs to add the residues between them to the QM region).

In QSite 4.0, QSite cuts are parametrized to the OPLS\_2001 force field. Side-chain and backbone cuts can be made as described in this section only if the force field option selected in the QSite Potential folder is OPLS\_2001, the default.

Cuts in a protein-ligand complex must be between atoms in peptide residues. Covalently-bound ligands can be included in the QM region, but only along with attached protein atoms. The QM region must extend at least as far as the first permissible cut between protein atoms.



**Figure 5.6.** The QM Settings *folder of the QSite panel.*

The QM Settings folder features are described below.

### Method

The options for the QM method are Density Functional Theory (DFT-B3LYP), Hartree-Fock, and Local Møller-Plessett perturbation theory (Local MP2).

### Charge

This is the net charge of the QM region of the system. Maestro updates the Charge to a reasonable value whenever a new residue or ion is added to the QM region; if a discrepancy appears, edit the value. If this value does not match the sum of the formal charges of the atoms in the QM region, Maestro displays a warning message, but allows you to proceed.

### Multiplicity

Check that this is the associated spin multiplicity of the QM region of the system: 1 for singlet, 2 for doublet, etc. Edit the value if necessary. If there is a discrepancy between the total charge and the multiplicity, the Jaguar calculation will halt with an error message. The charge and multiplicity of the QM region must be mutually consistent. By default, open-shell calculations

are spin-restricted. If you want to perform a spin-unrestricted calculation, you can add the keyword setting `iuhf=1` in the QM options text box.

### QM options

This text box can contain any Jaguar keywords such as print settings, non-default convergence criteria, and so on. Each such option is of the form `keyword=value` (with no embedded blanks). Multiple keyword/value pairs can be specified, separated by one or more blank spaces. By default, the following QM options appear in the box:

```
iacc=1 vshift=1.0 maxit=100
```

You can remove or modify these options as appropriate. See [Chapter 9](#) of the *Jaguar User Manual* for more information on Jaguar keywords.

### Define QM region section

In this section you define the residues that belong to the QM region of the calculation and set any related options. Residues selected for the QM region are listed in the table at the top of this section, one per row. Maestro assigns a residue name and number to every part of the Workspace structure. In the context of QSite residue can be an amino acid residue, a free ligand or solvent molecule, or an ion. There are three column headings:

- Type of cut

The three types of cut are `sidechain`, `backbone`, and `ligand/ion`, corresponding to the three selection states on the Pick option menu.

- Residue/Ion Name

Cuts are listed with this column showing residue name and residue number. For backbone cuts, each residue selected for the QM region is listed. For other types of cut, the residue listed is the one where the cut was made.

- Basis Set

The basis set associated with each residue in the list is shown in this column.

### Pick option and menu

This option allows you to pick the residues in the Workspace to define the QM region. The options for picking in the QM Settings folder are the three types of selections that can be added to the QM region:

- Side chains

Choose this option to add only the side chain of an amino acid residue to the QM region, leaving the backbone in the MM region. Pick an atom in the side chain you want to

include. The side chain is marked in sienna. A cut is made between the alpha carbon and the beta carbon. All of the atoms in the side chain are part of the QM region.

Side-chain cuts can be made in any peptide residue other than alanine (ALA), glycine (GLY), and proline (PRO). To incorporate side chains from these residues in the QM region, you must select the entire residue by using a backbone cut.

**Note:** If you are using the OPLS1999 force field, side-chain cuts in threonine (THR), serine (SER), and arginine (ARG) are not supported. The QSite calculation will fail if you make these cuts.

- Entire Residue

To add entire amino acid residues to the QM region, choose this option and then pick two atoms that are not alpha carbons and are at least three backbone bonds apart. The residues containing these two atoms and the residues in between are included in the QM region. The cut between MM and QM atoms is made between the alpha carbon and one of the atoms bonded to it so that the picked atoms are included in the QM region. When you pick the first atom it is marked with a purple cube. After picking the second atom, all of the backbone and side chain atoms between the two cuts are marked in sienna to show that they have been selected for the QM region. Backbone cuts can be made in any peptide residue, including glycine, proline, and their adjacent residues. There is one exception: backbone cuts on PRO residues cannot be made between the N atom and the C $\alpha$  atom.

**Note:** If you are using the OPLS1999 force field, backbone cuts for glycine (GLY), proline (PRO), and their adjacent residues are not supported. The QSite calculation will fail if you make these cuts.

- Ligand/Ions

Free ligands, metal ions, or other species not covalently bound to the protein can be added to the QM region by this method, which does not make any cuts between atoms. Select this option, then pick a metal ion or an atom in the ligand molecule to add it to the QM region. Molecules are marked in sienna. Single atoms or ions are marked in green.

Ligands that are covalently bound to the protein cannot be added using this method, because this method does not make parametrized cuts between bonded atoms. To add covalently-bound ligands to the QM region, make either a pair of backbone cuts to select the residue to which the ligand is bound, or make a side-chain cut.

#### Basis set option menu

By default, the basis set used for the entire QM region is 6-31G\* (LACVP\* in the case of transition metals), which is the basis set used in parametrization. To use a different basis set for any species in the QM region list, select its row, then choose another basis set from this menu.

**Note:** The OPLS2001 force field enforces 5D for the d functions; the OPLS1999 force field enforces 6D for the d functions.

#### Show markers check box

If Show markers is selected, a red trace marks the residues selected for the QM region. Any residue that is part of the QM region will also have its chain name, molecule number, residue number, and insertion code (if applicable) included in the list at the top of the panel.

#### Delete and Delete All buttons

To remove a specific residue from the QM region, select the residue in the QM region list and click Delete. When deleting individual residues, take care to ensure that the resulting residue set is consistent with any backbone cuts that have been made. To remove all residues from the QM region list, click Delete All.

## 5.7 Running QSite Jobs

When you have set the options in the QSite panel folders to the desired value, click the Start button to open the Start dialog box. In this dialog box you can set up the job options, then submit the job to a host for execution.

In the Output section, you can choose how the results of the job are incorporated into the project. The Incorporate option menu offers three options:

- Append new entries—append the results to the Project Table as new entries.
- Replace existing entries—replace the entry used as input with the resulting structure.
- Do not incorporate—Do not incorporate the results into the Project Table.

The Job section allows you to specify options for the execution of the job:

- Name text box—enter the job name. This is used to name the input and output files. The default job name is `qsites`.

**Note:** QSite does *not* automatically assign new names to jobs or files. If files of the same name exist, a warning is displayed before any files are overwritten.

- Host option menu—choose a host on which to run the job. This option menu displays all the hosts defined in the `$SCHRODINGER/schrodinger.hosts` file, with the number of processors on the host in parentheses. The default is `localhost`.
- Username text box—enter a user name if your username on the chosen host differs from your username on the host on which you are running Maestro.
- CPUs text box—enter the number of CPUs to use for parallel processing of the QM part of the calculation, if the host has multiple processors. You must have a parallel Jaguar license to run on multiple CPUs.

**Note:** Do not use parallel processing for jobs that include continuum solvation treatment, as they will fail if you do.

- Scratch directory option menu—select a scratch directory for storage of temporary files.

When you have selected job options, click the **Start** button to run the job. QSite jobs can be monitored in the job **Monitor** panel.

If a QSite job needs to be restarted, it must be restarted from the command line. To restart a job, simply run QSite with the restart file specified as the input:

```
$SCHRODINGER/qsites job-options jobname.01
```

The restart input file has an index number inserted into the file name: for the first restart it is `jobname.01.in`. Running with this file produces a restart file named `jobname.02.in`, and so on.

If you want to run calculations on systems with more than 8000 atoms or 8000 bonds, you must run them from the command line, but you can still set up the job from the QSite panel.

#### To run a large QSite job:

1. Set up the job in the QSite panel.
2. Click **Write**.
3. Edit the `jobname.in` file and add the keyword setting `impversion=huge` to the **gen** section.
4. Start the job from the command line with the standard QSite syntax:

```
$SCHRODINGER/qsites job-options jobname
```



# Running QSite from the Command Line

Although you will normally set up QSite jobs using the controls and settings in the Maestro GUI, you can submit jobs either from Maestro or from the command line. The same is true for the Protein Preparation facility. Advantages of running from the command line include:

- The command-line scripts can run all full-featured jobs written using the QSite and Impact panels in Maestro, and also allow you to override specific run-time values that are not accessible through the Maestro interface.
- Command-line scripts allow you to run jobs when you want.
- Command-line scripts can be modified and jobs can be re-run without reconfiguring and reloading job settings in Maestro.
- Some job options are available only when you run QSite from the command line.

The `SCHRODINGER` environment variable must be set to run jobs. You can define `SCHRODINGER` as follows:

**csh/tcsh:**        `setenv SCHRODINGER installation-directory`

**bash/ksh:**        `export SCHRODINGER=installation-directory`

Unless otherwise specified, Schrödinger applications and utilities run under the Schrödinger Job Control facility and are automatically run in the background. For more information on the Job Control facility, see the *Job Control Guide*.

## 6.1 QSite Files

QSite jobs have a single input file, which is like the Jaguar input file. The input for the QM part of the calculation is included as regular Jaguar sections with their keywords. The input for the MM part is embedded in an **impact** section, and consists of the Impact commands.

The Write button in the QSite panel writes the input files needed for a job. See [Section 5.1.2 on page 58](#) for more information.

Like Jaguar, QSite writes a restart file at various stages of the calculation. The restart input file has an index number inserted into the file name: for the first restart it is `jobname.01.in`. Running QSite with this file produces a restart file named `jobname.02.in`, and so on.

A typical QSite job has one input (*jobname.in*), one or more structure files (*jobname.mae*, *jobname.pdb*, or *jobname.sdf*), and after execution, several output files (e.g., *jobname\_out.mae* for structure files and *jobname.out* for textual data).

If a file already has the name of an output file, in many cases Impact renames the old file with a numerical extension (*filename.out.01*, *filename.out.02*, and so on) for archival purposes. The new job's output is then written to the base name (*filename.out*). If you do not need the old files, you can remove them.

Some files, such as *jobname.log* files, are newly written each time Impact runs a calculation. Other examples of files that are *not* incremented are *jobname\_out.mae* structure files, for Basic Impact minimization and QSite jobs.

Table 6.1 contains descriptions of the various file types. For more information, see the Maestro online help or the *Impact Command Reference Manual*.

Table 6.1. QSite and Impact file extensions

Extension	Description
.mae	A Maestro format structure file, a plain-text file written by Maestro containing atom, bond, and other information for one or more molecules.
.log	An Impact log file. If specified, a .log file captures standard output and standard error messages in text form. This file is overwritten during subsequent runs.
.in	The input file for a QSite calculation. Includes both Jaguar and Impact input.
.out	An Impact output file containing information similar to that found in log files (no standard error). Output files are appended with numerical extensions when the input file is run again. Up to 99 output files are retained.
.01, .02, etc.	A file containing results from previous Impact calculations run from the corresponding <i>jobname.inp</i> file.
_out.mae	An Impact output structure file written in the Maestro file format. Some Impact jobs do not write *_out.mae output structure files.

## 6.2 The qsite Command

The `qsites` command line provides both the means to submit QSite jobs and to interact with the Job Control facility to query and control the progress of a job. To run a job, the syntax is:

```
$SCHRODINGER/qsites [run-options] jobname[.in]
```

The `.in` suffix for the input file, *jobname.in*, is optional: if it is not specified, it is added. The options for running QSite calculations include the standard Job Control options, and are described in Table 6.2.

Table 6.2. Options for running QSite jobs

Option	Description
-ARCH <i>platform</i>	If you have more than one architecture installed for a given system, e.g., AIX-com and AIX-pwr3, then this option can be used to select either of them, such as -ARCH pwr3.
-DEBUG	print detailed information about job launch to terminal
-HOST <i>hostname</i>	Run the job on the host <i>hostname</i>
-PROCS <i>nprocs</i>	Run the job using <i>nprocs</i> processors. A parallel Jaguar license is required for this option. Do not specify multiple processors if continuum solvation is in use.
-REL <i>release</i>	This option selects a specific version of QSite to use. The default is the latest (highest number). Formats like -REL v4.0, -REL 35000, and -REL 27 are supported.
-SAVE	Save the scratch directory (do not delete it at the end of the job).
-USER <i>user</i>	Specify a user name on the host specified by -HOST. Default is to use the user name from the local host.
-VER <i>pattern</i>	If you have multiple versions installed, you can specify a pattern with this option that matches the installation path to use for your job. The default installation is the one printed by -WHICH.
-WAIT	Wait for the job to finish before executing the next command. If the job is run from a terminal window, the prompt is not displayed until the job finishes. This is useful in command scripts in which you have specified actions to take only after the QSite job finishes.

You can use the `qsite` job to query Job Control for information about hosts and software versions. The syntax of the query command is

```
$SCHRODINGER/qsite query-option
```

The available information options are listed in [Table 6.3](#).

The `qsite` command also has its own commands for querying Job Control about jobs, controlling jobs, and obtaining some other information. The syntax is as follows:

```
$SCHRODINGER/qsite command command-options
```

The available commands are listed in [Table 6.4](#). Further information is available in the *Job Control Guide*.

Table 6.3. Information options for the *qsite* command.

Option	Description
-WHICH	List the available installations for the local machine. The job itself is not submitted. The first one listed is the default path; the options -REL, -VER, and -ARCH can direct your job to use a different installation.
-HOSTS	List the available hosts.
-ENTRY	Show the host file ( <code>schrodinger.hosts</code> ) entry for the host specified with the -HOST option.
-LIST [-ALL]	List the available compatible versions; list all versions, including incompatible versions, if -ALL is used.

Table 6.4. Syntax for *qsite* commands.

Command	Description
<code>jobs [name   status]</code>	Show status of the specified QSite jobs or list the jobs that have the specified status.
<code>kill {name   status}</code>	Kill the QSite jobs with the specified job name or job status.
<code>stop {name   status}</code>	Stop the QSite jobs with the specified job name or job status at the end of the current calculation stage.
<code>purge [name]</code>	Remove records for the specified jobs from the job database. If no <i>name</i> is given, all completed jobs are purged.
<code>help</code>	Display usage message
<code>hosts</code>	
<code>config</code>	
<code>machid</code>	Report the hardware and software configuration. This command gives the same output as the <code>\$SCHRODINGER/machid</code> command.
<code>platform</code>	Report information on the hardware platform. This command gives the same output as the <code>\$SCHRODINGER/platform</code> command.
<code>sysreq</code>	Report any system requirements for QSite and whether they are met.

There are two Impact executables, compiled with different atom limits. Impact guesses which executable to use based on molecular size values given in the command input file, but you can also select the executable with the `-s` option. should be used on systems of greater than 8000 atoms or 8000 bonds.

## 6.3 Command-Line Applications and Utilities

The command-line application `protprep` is located in the main Schrödinger directory. Utilities are located in the directory `$SCHRODINGER/utilities`. You may want to add this directory to your path so that they are easy to run by name from the command line. For usage summary information, use the `-h` (help) option:

```
$SCHRODINGER/protprep -h
$SCHRODINGER/utilities/utilityname -h
```

Protein preparation jobs can be run from the command line using the `protprep` application:

```
$SCHRODINGER/protprep [options] input-file
```

You can also use the `pprep` and `impref` utilities. See [Section 4.11 on page 60](#) for more information about command-line protein preparation.

A utility is also provided to convert separate Impact and Jaguar input files from releases prior to 4.0 into a single QSite input file. The syntax is:

```
$SCHRODINGER/utilities/mkqsinput jobname.inp
```

where `jobname.inp` is the name of the Impact input file. The corresponding Jaguar input file, `jobname.jaguar.in`, is assumed to be in the same directory as the `jobname.inp` file. The output is a file named `jobname.in`. This utility changes OPLS1999 to OPLS2001, because the use of OPLS\_2001 is strongly recommended over OPLS\_1999. If you wish to retain the OPLS\_1999 force field, edit the new input file and change the force field.



# QSite Technical Notes

The study of reactive chemistry in a protein environment is an extremely challenging problem for computational chemistry. The only methods that can produce reliable results (particularly for structures containing transition metals) are those of *ab initio* quantum chemistry. However, such methods are computationally intensive and scale poorly ( $\sim N^2$ – $N^3$  for self-consistent-field based approaches such as density functional theory) with the number of atoms  $N$ , making it impractical to apply them to an entire protein. An attractive solution is use a mixed quantum mechanics/classical molecular mechanics (QM/MM) method such as QSite. Such methods treat the reactive core of the system quantum mechanically and model the remainder via a classical molecular-mechanics force field.

## 7.1 QM/MM for Protein Active Sites

QSite is specifically designed to treat protein active sites. It combines Schrödinger's powerful Jaguar program for *ab initio* electronic structure calculation with molecular-mechanics calculations that use the OPLS-AA force field of Jorgensen and coworkers. The speed of Jaguar—augmented by an MPI-based parallel implementation—makes it possible to study realistic representations of the active site with the large QM regions (typically 100–200 atoms in applications we have pursued) that are in many cases necessary to obtain chemically realistic results. Hartree-Fock (HF), density functional (DFT), and local MP2 (LMP2) methods are available for the QM/MM region, although geometry optimization has been implemented only for the first two of these methods. We have found DFT methods to be particularly useful for studying protein active-site reactive chemistry.

QSite provides QM/MM interface parameters for all 20 amino acids in their various protonation states. This ensures that you will be able to construct a QM region tailored specifically to your needs. Maestro provides an easy and reliable way to set up the QM/MM interface via pointing and clicking with the mouse on residues of the protein active site. Protein preparation can be carried out using the procedure described in [Chapter 4](#). This technology makes setting up a new QM/MM job a task that can be carried out effectively and straightforwardly.

One key use of QSite is to study ligand binding to transition-metal-containing enzymes such as zinc matrix metalloproteases. Conventional molecular-mechanics force fields usually model ligand interactions with protein metals in a primitive fashion, i.e., as a van der Waals body and a charge site. Jaguar contains specialized methods for treating transition metals that include a novel initial-guess methodology and variable-energy-shift algorithms to converge difficult

cases. The combination of Jaguar's ability to handle large systems efficiently and to accelerate the SCF convergence yields a methodology of unprecedented power and flexibility. No other commercially available program can provide the kind of chemical insight and quantitative description for metal-containing enzymes that QSite offers. As described below, QSite also allows reactive processes to be modeled and transition states to be located. For applications such as these, QSite is an essential part of a comprehensive computational strategy for structure-based drug design.

## 7.2 QM/MM Transition State Modeling

Jaguar and QSite can perform transition state (TS) searches by using a quasi-Newton method to find the TS nearest the initial geometry [1]. Alternatively, Jaguar can employ a Linear/Quadratic Synchronous Transit (LST/QST) approach, which is also known as Synchronous Transit Quasi-Newton (STQN), to guide the search along the reaction pathway between specified reactant and product geometries [2]. This latter approach is clearly superior and has now been extended for use in QSite.

Through the Maestro GUI, you can enter one of the following:

1. Initial guess for the TS (a simple TS search is then used)
2. Reactant, product, and TS guess (QST is used, with the entered guess used as the initial TS geometry)
3. Reactant and product geometries (LST is used, but an initial TS geometry is generated automatically)

For case 3, the automatic generation is done by interpolating between the reactant and product structures for only the atoms seen by the QM program, Jaguar (these atoms include those in the QM region plus a small number of MM atoms located at the QM/MM interface). This procedure includes a least-squares fit of the interfacial MM atoms in the interpolated geometry to the respective atoms in each of the reactant and product geometries. The interpolated TS QM region plus interfacial MM atoms are then inserted into the pure MM structure of the best-fitting case (reactant or product) using the transformation found from the least-squares fitting.

The LST/QST guided search (for cases 2 or 3) then proceeds as it does for Jaguar by first restricting the optimizer to search along the circular curve connecting the reactant, TS, and product structures. Again, only the QM plus interfacial MM atoms seen explicitly by Jaguar are used. The pure MM atoms are adiabatically minimized at each step. Once the optimizer approaches (or finds) a maximum-energy TS structure along this reactant-product curve, the TS search proceeds along the Hessian eigenvector that is most similar to the tangent to the circular curve. This process continues until a saddle point with one negative eigenvalue (corresponding to an imaginary frequency) is found.

In contrast, a simple transition state search (case 1) just involves the attempt by the optimizer to maximize the energy along the lowest-frequency eigenvector of the Hessian and to minimize along all other modes. Again, only the QM and interfacial MM atoms are included in the determination of the Hessian, as all pure MM atoms are adiabatically minimized at each TS search step.

## 7.3 How QSite Works

Most approaches for developing robust and accurate QM/MM methods have been based on “link atom” approaches, in which QM and MM fragments are capped by hydrogens. These methods face nontrivial problems in constructing an accurate description of the QM/MM interface, particularly for polar systems, where the treatment of electrostatic interactions can be highly problematic. While progress has been made, we do not believe that a fully satisfactory link-atom methodology is available. QSite takes an alternative approach in which frozen localized molecular orbitals are used to build the QM/MM interface. This methodology has recently been reviewed favorably [3]. As far as we are aware, QSite is the first ab initio frozen-orbital methodology with analytical gradients for which accuracy for structures and conformational energetics of a polar system has been demonstrated.

The details of QSite’s frozen-orbital interface technology is provided in References 4-6 at the end of this chapter. The key aspects are:

- The frozen orbital itself is obtained by Boys localization of the quantum chemical wavefunction for one of a series of small template molecules. The orbital must be translated and rotated as the molecular geometry changes, and its interaction with both the QM and MM regions must be properly represented. The charge distribution must be empirically corrected to reproduce the fully quantum chemical result. QSite does this by placing a charge in the middle of the frozen bond. QSite not only includes appropriate energy expressions for this representation but also the analytical gradients that are critical to applications that involve geometry optimization.
- The QM and MM regions interact via two mechanisms: Coulomb interaction between MM charges and the QM wavefunction, and van der Waals interaction between QM and MM atoms (both of which employ van der Waals parameters).
- Specialized correction terms are used for stretches, bends, and torsions involving the atoms directly associated with the frozen-orbital interface. These terms are fit to reproduce quantum-chemical conformational energies of the template molecules. Again, QSite has gradients for all of these terms as well as energy expressions.

The torsional correction parameters were determined from a library of high-quality QM calculations on rotamer states for dipeptides. Beginning with roughly 300 geometries obtained via

conformational search using OPLS-AA, the structures were optimized at the HF/6-31G\*\* level and single point LMP2/cc-pVTZ (-f) relative energies were computed. Finally, one-dimensional torsional profiles were generated at the same level of theory for all minima and relevant torsional degrees of freedom (~2000 QM data points in all). Alanine tetrapeptide conformations, generated via the same protocol, were used to test transferability. In addition, a database consisting of hydrogen-bonded pairs of small-molecule side-chain analogues was constructed. About 200 such pairs were used to determine van der Waals radii for QM atoms that yield accurate hydrogen bonding energetics between QM and MM donors and acceptors. These data sets are considerably more extensive, and of higher quality, than any that have been used previously in developing or testing QM/MM models of peptides and proteins.

Both DFT and Hartree-Fock parameter sets have been developed. The LMP2 version of the theory has been implemented for use in “single-point” calculations. It has not yet been fully parameterized, but can be used to compare structures and energies when there is little change in the protein geometry in the vicinity of the frozen orbital interfaces. Details of the parameter optimization methodology are provided in the previously cited references.

QSite makes use of a tight coupling between Jaguar and Impact. Key features of the implementation are as follows:

- QSite adiabatically minimizes the MM region after each QM geometry step. Without this, the number of QM steps would become prohibitively large and would place the calculations out of range for all but the most powerful supercomputers.
- QSite can run the rate-limiting QM part of the code on parallel processors so that reasonable throughput can be achieved for the relatively large (100–200 atom) QM regions that can be necessary to reliably model active-site reactive chemistry.
- QSite incorporates a Poisson-Boltzmann continuum-dielectric treatment of aqueous solvation. This treatment is capable of handling the QM and MM regions simultaneously and includes an analytic gradient, so that geometry optimization in solution can be performed. Inclusion of solvation will be critical in some (but not all) applications, an obvious case being calculation of  $pK_a$  values of ionizable protein side chains or of ligand groups that interact with the protein.

Maestro makes it easy to set up a QSite calculation. For example, its QSite interface can be used to readily identify the QM residues of the protein via mouse clicks, and to specify whether the QM/MM “cut” is to be placed in the backbone or side chain. The 6-31G\* basis set is used in the interface region, but other basis sets can be used elsewhere. For example, the geometry can be optimized with the 6-31G\* basis set, and a large, high quality basis set can then be used in a reactive region to determine accurate single-point energies. We have used this strategy very successfully in our QM-based modeling of the protein methane monooxygenase [7] and expect it to work for a wide range of active-site modeling applications.

## 7.4 Parametrization Validation

### 7.4.1 Deprotonation Energies

Chemical reaction energetics provide one important measure of how well a QM/MM model reproduces accurate quantum mechanics. Below we examine a simple reaction—removal of a proton from the QM region. [Table 7.1](#) compares differences between the QM and QM/MM deprotonation energies for the capped peptides we have examined and lists the distance between the proton to be removed and the frozen interface orbital. When this distance is greater than  $\sim 5 \text{ \AA}$  (a distance that one would want to maintain between a reactive chemical event and the QM/MM interface in any event) the error is less than 0.4 kcal/mole. This error is negligible when compared to total reaction energies of hundreds of kcal/mole and is small even in comparison to the intrinsic error HF or DFT calculations make relative to experiment. (DFT does quite well with a large basis set, but still makes errors on the order of 1–2 kcal/mole for small-molecule deprotonation energies). When the reactive event is very close to the frozen bond, the errors can be somewhat larger but are still very reasonable.

*Table 7.1. B3-LYP/6-31G\* QM/MM absolute deprotonation energy differences relative to fully QM B3-LYP/6-31G\* values. (\*) denotes the deprotonated QM leucine residue (H of C $\delta$ ). The line through the capitalized residue denotes the QM/MM boundary. This residue is mostly in the QM region and the cut is made between N and C $\alpha$ .*

Peptide MM Region	Peptide QM Region	Frozen orb.-H dist. (Å)	Error (kcal/ mole)
ace - ala - <b>L</b>	EU* - nma H1	4.1	0.70
ace - ala - <b>L</b>	EU* - nma H2	4.3	0.52
ace - ala - <b>L</b>	EU* - nma H3	4.3	0.53
ace - ala - <b>A</b>	LA - leu* - nma H1	5.2	0.40
ace - ala - <b>A</b>	LA - leu* - nma H2	5.5	0.34
ace - ala - <b>A</b>	LA - leu* - nma H3	6.3	0.29
ace - ala - <b>A</b>	LA - ala - leu* - nma H1	8.7	0.20
ace - ala - <b>A</b>	LA - ala - leu* - nma H2	9.7	0.15
ace - ala - <b>A</b>	LA - ala - leu* - nma H3	8.2	0.23
ace - ala - ala - ala - ala - <b>L</b>	EU* -ala - nma H1	4.1	1.21
ace - ala - ala - ala - ala - <b>L</b>	EU* -ala - nma H1	4.3	0.93
ace - ala - ala - ala - ala - <b>L</b>	EU* -ala - nma H1	4.3	0.82

### 7.4.2 Conformational Energies

Conformational energies afford a second critical test of the quality of the QM/MM methodology. Table 7.2 summarizes the RMS deviations between the QM/MM or MM (OPLS-AA) and the purely QM results for the relative energetics of the side-chain rotamer data set used to parameterize QSite. The QM/MM results generally are at least as good as those from the OPLS-AA force field, and in many cases are much better. The QM/MM rotamer structures also display good fidelity to the QM structures.

### 7.4.3 Other Comparisons

QSite has also been shown to accurately reproduce patterns in quantum-mechanical hydrogen-bond dimerization energies and to reproduce LMP2/cc-pVTZ(-f) relative conformational energies for the alanine dipeptide more accurately than does a molecular force field such as OPLS-AA (see Ref. 6).

Table 7.2. RMS deviations (kcal/mole) of rotamer side chain conformational energies for HF/6-31G\* QM-MM, B3LYP/6-31G\* QM-MM and OPLS-AA relative to cc-pVTZ(-f) LMP2

Residue	QM-MM (QM=HF)	QM-MM (QM=DFT)	OPLS-AA
Phe	0.18	0.14	0.18
His	1.14	1.10	1.05
Asn	0.38	1.36	2.29
Val	0.21	0.41	0.62
Trp	0.71	0.25	0.75
Tyr	0.37	0.29	0.40
Leu	0.77	0.90	0.40
Met	1.10	1.21	1.82
Gln	1.48	1.03	2.70
Glu	2.61	2.69	3.34
Cys	0.61	0.62	3.52
Lys	1.44	1.29	4.22
Ile	1.03	0.56	1.19
Asp	1.11	1.91	2.51
Arg	1.67	3.50	2.90

## 7.5 An Illustrative Application

Cytochrome P-450 is an enzyme whose variants are ubiquitously distributed across a wide variety of organisms. A human version of the enzyme in the liver is of great importance pharmaceutically because it is involved in a significant fraction of toxicity and drug metabolic pathways. While a high-resolution structure of a form of the human enzyme relevant to questions of toxicology and metabolism does not yet exist, such structures are likely to be produced in the next few years (either experimentally or via homology modeling). This will open the possibility of computer modeling of these critical processes.

It is very difficult to study the chemistry of ligand binding of cytochrome P450 with conventional molecular-modeling techniques. The existence of a reactive metal center, and the centrality of the reactive chemistry in the interaction of the enzyme with various drug candidates, mandate a quantum-chemical treatment of the active site. On the other hand, the protein structure is clearly important in selectivity, binding affinity, and reaction kinetics, and cannot be incorporated to any great degree with conventional quantum-chemical techniques. QSite is well suited to addressing this problem with relatively modest computational resources. Its application to a problem in cytochrome P450 chemistry [6] is recounted below.

### 1. Structure and QM Region

The particular systems studied are available as entries 1phf and 1akd in the PDB archives. The former has a coordinating phenylimidazole ligand, while the latter is the so-called substrate-free state with a non-coordinating camphor ligand located near the heme.

### 2. P-450 and Phenylimidazole Ligand

The 1phf structure has a 4-phenylimidazole complexed with Fe. On the opposite side of the heme, Fe is coordinated to a cysteinate sulfur ( $\text{R-S}^-$ ). Thus, the Fe is in a six-fold coordination site and is assumed to be  $\text{Fe}^{3+}$ , given the  $\text{S}^-$  formal charge and the two negative charges distributed over the four coordinating nitrogens in the heme ring system. The net spin state is either a doublet, with the Fe-S moiety low-spin coupled as in the coordinated dioxygen state of P-450, or a high-spin quartet. There are 7075 atoms in the system.

The QM region is the full heme ring, the Fe, the full coordinating cysteine residue including residues on either side of the cysteine, and the 4-phenylimidazole. The two QM-MM cuts were made in the residues adjacent to the coordinating cysteine. The net charge of the QM region is  $-2$  from the two carboxylate groups on the heme. There are 125 total quantum atoms and 1138 6-31G\* basis functions. The QM method used was B3LYP DFT. The outer shell of the protein was frozen during the optimization, leaving 3960 of the 7075 atoms free to optimize. This procedure is commonly used to avoid irrelevant energy differences caused by re-arrangements of the outer parts of the protein.

### 3. P-450 and Camphor

The specification of this system (1akd) is identical to that above, but with a quantum camphor molecule replacing the phenylimidazole ligand. The camphor is not directly bound to the Fe but, like the phenylimidazole ligand, has forced water out of the heme-binding region. The main purpose of running this system was to find the lowest spin state of the substrate-free P-450.

### 4. Preliminary results

The P-450 optimization was run on 6 SGI-R10000 nodes in about a 1 week of wall-clock time. The calculation took ~30 QM-MM geometry-optimization cycles. Approximately 10 hours was spent on the initial MM minimization in which the QM region is frozen. The final geometry has an RMS deviation of 0.6 Å with respect to the non-hydrogen atoms of the crystal structure, a reasonable level of accuracy. In a general study of a system like this, the initial cost of this minimization would be amortized over similar runs in which the QM region would be perturbed (by changing the ligand, for example) since these subsequent runs would have a good initial geometry.

As an initial calibration of the energetics we optimized the doublet and quartet spin states of the system. The doublet was found to be 14 kcal/mole lower in energy. This is in accord with qualitative EPR data [8] that indicates that Fe is in a low-spin state when a six-fold  $\text{Fe}^{3+}$  coordination site involves dative bonding to the ligand. The substrate-free system with the camphor above the heme ring was found after full QM-MM optimization to have a high-spin quartet ground state with the doublet 15 kcal/mole higher in energy. This ordering of spin states is also in agreement the experimental ordering [9].

## 7.6 References

1. Bofill, J. M. *J. Comp. Chem.* **1994**, *15*, 1; Murtagh, B. A.; Sargent, R. W. H. *Comput. J.* **1970**, *13*, 185; Powell, M. J. D. *Math. Prog.* **1971**, *1*, 26.
2. Peng, C.; Schlegel, H. B. *Isr. J. Chem.* **1993**, *33*, 449.
3. Gogonea, V.; Suárez, D.; van der Vaart, A.; Merz, K. M. Jr. *Curr. Opin. Struct. Biol.* **2001**, *11*, 217–223.
4. Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **1999**, *20*, 1468.
5. Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *Chem. Phys. Lett.* **2000**, *321*, 113.
6. Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442.

7. Dunietz, B. D.; Beachy, M. D.; Cao, Y.; Whittington, D. A.; Lippard, S. J.; Friesner, R. A. Large scale ab initio quantum chemical calculations of intermediates in the soluble methane monooxygenase catalytic cycle. *J. Am. Chem. Soc.* **2000**, *122*, 2828–2839.
8. Lippard, S. J.; Berg, J. M. *Principles of Bioinorganic Chemistry*. University Science Books, 1994.
9. Porter, T. D.; Coon, M. J. Cytochrome P450—multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. *J. Biol. Chem.* **1991**, *266*, 13469–13477.



# Getting Help

Schrödinger software is distributed with documentation in PDF format. If the documentation is not installed in `$SCHRODINGER/docs` on a computer that you have access to, you should install it or ask your system administrator to install it.

For help installing and setting up licenses for Schrödinger software and installing documentation, see the *Installation Guide*. For information on running jobs, see the *Job Control Guide*.

Maestro has automatic, context-sensitive help (Auto-Help and Balloon Help, or tooltips), and an online help system. To get help, follow the steps below.

- Check the Auto-Help text box, which is located at the foot of the main window. If help is available for the task you are performing, it is automatically displayed there. Auto-Help contains a single line of information. For more detailed information, use the online help.
- If you want information about a GUI element, such as a button or option, there may be Balloon Help for the item. Pause the cursor over the element. If the Balloon Help does not appear, check that Show Balloon Help is selected in the Help menu of the main window. If there is Balloon Help for the element, it appears within a few seconds.
- For information about a panel or the folder that is displayed in a panel, click the Help button in the panel. The Help panel is opened and a relevant help topic is displayed.
- For other information in the online help, open the Help panel and locate the topic by searching or by category. You can open the Help panel by choosing Help from the Help menu on the main menu bar or by pressing CTRL+H.

To view a list of all available QSite-related help topics, choose QSite from the Categories menu of the Categories tab. Double-click a topic title to view the topic.

If you do not find the information you need in the Maestro help system, check the following sources:

- *Maestro User Manual*, for detailed information on using Maestro
- *Maestro Command Reference Manual* for information on Maestro commands
- *Job Control Guide*, for information on running jobs
- *Jaguar User Manual*, for information on Jaguar and its keywords
- *Impact Command Reference Manual*, for information on QSite and Impact commands
- *Impact User Manual*, for information on running Basic Impact applications
- Frequently Asked Questions pages, at [https://www.schrodinger.com/QSite\\_FAQ.html](https://www.schrodinger.com/QSite_FAQ.html)

The manuals are also available in PDF format from the Schrödinger [Support Center](#). Information on additions and corrections to the manuals is available from this web page.

If you have questions that are not answered from any of the above sources, contact Schrödinger using the information below.

E-mail: [help@schrodinger.com](mailto:help@schrodinger.com)

USPS: Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204

Phone: (503) 299-1150

Fax: (503) 299-4532

WWW: <http://www.schrodinger.com>

FTP: <ftp://ftp.schrodinger.com>

Generally, e-mail correspondence is best because you can send machine output, if necessary. When sending e-mail messages, please include the following information, most of which can be obtained by entering `$SCHRODINGER/machid` at a command prompt:

- All relevant user input and machine output
- QSite purchaser (company, research institution, or individual)
- Primary QSite user
- Computer platform type
- Operating system with version number
- QSite version number
- Maestro version number
- mmshare version number

---

# Copyright Notices

## **C and C++ Libraries for Parsing PDB Records**

The C and C++ libraries for parsing PDB records are Copyright (C) 1989 The Regents of the University of California. All rights reserved.

Redistribution and use in source and binary forms are permitted provided that the above copyright notice and this paragraph are duplicated in all such forms and that any documentation, advertising materials, and other materials related to such distribution and use acknowledge that the software was developed by the University of California, San Francisco. The name of the University may not be used to endorse or promote products derived from this software without specific prior written permission. THIS SOFTWARE IS PROVIDED “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.



## A

algorithms	
Impact energy minimization .....	65
transition state search .....	67
atom constraints .....	62
atom types, correcting .....	46, 48
atoms	
constrained .....	62
harmonically restrained .....	62
selecting .....	21–23
Auto-Help .....	28, 91

## B

backbone, leaving in MM region .....	70
Balloon Help .....	28, 91
basis sets .....	72
bonds, protein-metal .....	45
Build panel .....	19
building structures .....	18–21
button menu .....	7

## C

Cartesian coordinates	
constraints .....	62
optimization of .....	66
Command Script Editor panel .....	24
command scripts— <i>see</i> scripts	
constant dielectric .....	59
constrained atoms .....	62
constraints	
Cartesian .....	62
frozen-atom .....	62
Constraints folder .....	62
continuum solvation .....	61
conventions, document .....	vii
convergence criteria, energy minimization .....	65
covalently bound ligands .....	71
current working directory .....	4
cuts, QSite	
definition .....	68
types .....	70

## D

dielectric constant .....	59
---------------------------	----

## directory

current working .....	4, 25
output .....	25
utilities .....	79
distance-dependent dielectric .....	60
duplicate binding sites .....	43

## E

electrostatic treatment .....	59, 60
energy minimization .....	66
algorithms .....	65
convergence criteria .....	65
cycles, maximum .....	65
steric clashes in .....	65
entries, Project Table .....	11
including, excluding, and fixing .....	16
merging for QSite .....	57
selecting .....	15
sorting .....	13
environment variables	
DISPLAY .....	4
SCHRODINGER .....	3–4, 29
ePlayer .....	13, 14
excluded entries .....	16

## F

file I/O directory .....	25
filters, project entry .....	15
fixed entries .....	16
force field .....	59
fragments, building structures from .....	18
frozen atoms .....	62
full screen mode .....	6, 11
function key macros— <i>see</i> scripts	

## G

geometry optimization .....	66
grow bond .....	19

## H

help option (-h) .....	79
Help panel .....	28, 91

**I**

impref utility .....	55
included entries .....	16
input files.....	76
conversion utility .....	79

**J**

Jaguar .....	57
keywords.....	70
specifying calculation.....	66
jobs	
restarting .....	73
running in Maestro .....	26–27

**K**

keywords, Jaguar.....	70
-----------------------	----

**L**

ligands	
adding to the QM region.....	71
covalently bound.....	71
protein preparation.....	48
Linear Synchronous Transit (LST) method .....	67
log file, saving Maestro.....	28
long range forces, in TN optimizations.....	66

**M**

macros— <i>see</i> scripts	
Maestro	
introduction to .....	1
main window .....	4, 5
menus.....	6
quitting.....	28
running jobs from .....	26–27
scratch projects .....	11
starting .....	4
undoing operations .....	26
main window .....	5
menu button .....	7
merging entries for QSite .....	57
metal ions	
adding to the QM region.....	71
adjusting atom type and charge .....	46
in protein preparation .....	40
Minimization folder .....	64

**MM region**

boundaries.....	68
constraints in.....	62
interaction with QM region .....	83
leaving backbone in .....	70
Monitor panel.....	27
mouse functions .....	3
Project Table panel .....	16–17
Workspace .....	10
multimeric protein structure.....	40, 42, 43

**N**

net charge, QM region .....	69
-----------------------------	----

**O**

online help.....	28
OPLS-AA force fields.....	59
Optimization folder .....	32, 67
output files.....	34, 76

**P**

parallel processing .....	1
restrictions .....	61
parametrization, QSite cuts	
basis set used .....	72
force field used .....	68
Potential folder.....	31
pprep utility .....	55
Preferences panel .....	25, 26
product installation.....	91
project entries, <i>see</i> entries, Project Table	
Project Facility, introduction.....	11
Project Table panel.....	13
menus.....	14
mouse functions .....	16–17
shortcut keys .....	17
projects .....	11
protein complex structure	
importing .....	40, 41
truncating .....	42
protein preparation, overview .....	40
protein, multimeric.....	42
protprep command-line application .....	53
Python scripts— <i>see</i> scripts	

**Q**

QM methods.....	69
QM options .....	70
QM region	
boundaries.....	68
defining .....	31, 70
frozen atoms .....	62
interaction with MM region.....	83
net charge.....	69
QM Settings folder.....	32, 69
QM/MM interface.....	83
QM/MM, defined .....	1
QSite	
brief description.....	1, 57
full description.....	81
QSite calculations	
available force fields .....	59
restrictions .....	57
unavailable solvation methods.....	61
QSite panel	
Minimization folder.....	64
Optimization folder .....	32, 67
Potential folder .....	31
QM Settings folder .....	32, 69
Quadratic Synchronous Transit (QST) method	68
quasi-Newton method .....	67
quitting Maestro .....	28

**R**

residues	
adding to the QM region.....	71
removing from the QM region.....	72
restarting jobs.....	73

**S**

Schrödinger contact information.....	92
scratch entries.....	12
scratch projects.....	11
scripts	
function key macros.....	25
macros.....	25
Maestro command .....	24

Python.....	23
utility.....	79
selecting objects in the Workspace .....	7, 21
shortcut keys	
main window .....	11
Project Table panel .....	17
side chains, adding to the QM region .....	70
single-point energy calculations .....	66
solvation, continuum.....	61
step size, Impact energy minimization.....	65
steric clashes, in energy minimization .....	65
structures	
building .....	18–21
displaying in sequence.....	13
importing .....	30

**T**

technical support .....	28
toolbar	
Build panel.....	20–21
main window .....	7–10
Project Table panel .....	13–14
transition-state optimization methods .....	67
Truncated Newton (TN) algorithm .....	66

**U**

undoing Maestro operations.....	26
utilities	
command-line .....	79
help option (-h) .....	79
impref .....	55
pprep .....	55

**W**

waters, deleting .....	42
Workspace	
description .....	4
full screen mode .....	6, 11
including, excluding, and fixing entries .....	16
mouse functions.....	10
scratch entries .....	12





---

120 West 45th Street  
32nd Floor  
New York, NY 10036

101 SW Main Street  
Suite 1300  
Portland, OR 97204

3655 Nobel Drive  
Suite 430  
San Diego, CA 92122

Dynamostraße 13  
68165 Mannheim  
Germany

QuatroHouse, Frimley Road  
Camberley GU16 7ER  
United Kingdom

**SCHRÖDINGER.**